PATENT COOPERATION TREA

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF THE RECORDING OF A CHANGE

(PCT Rule 92bis.1 and Administrative Instructions, Section 422)

To:

CHARI, Santosh, K.
Orange & Chari, Suite 4900, P.O. Box 190
66 Wellington Street West
Toronto Dominion Bank Tower
Toronto-Dominion Centre
Toronto, Ontario M5K 1H6
CANADA

Date of mailing (day/month/year)
02 November 2000 (02.11.00)

Applicant's or agent's file reference
00264-0003

IMPORTANT NOTIFICATION

International application No.
PCT/CA00/00430

International filing date (day/month/year)
20 April 2000 (20.04.00)

1. The following indications appeared on record concerning:

[ ] the applicant [ ] the inventor [X] the agent [ ] the common representative

Name and Address
CHARI, Santosh, K.
Orange & Chari
4900-55 King Street West, P.O. Box 190
Toronto Dominion Bank Tower
Toronto-Dominion Centre
Toronto, Ontario M5K 1H6
Canada

State of Nationality:
State of Residence:
Telephone No.: 416 601-8440
Facsimile No.: 416 601 8454
Teleprinter No.:

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

[ ] the person [ ] the name [X] the address [ ] the nationality [ ] the residence

Name and Address
CHARI, Santosh, K.
Orange & Chari, Suite 4900, P.O. Box 190
66 Wellington Street West
Toronto Dominion Bank Tower
Toronto-Dominion Centre
Toronto, Ontario M5K 1H6
Canada

State of Nationality:
State of Residence:
Telephone No.: 416 601-8440
Facsimile No.: 416 601 8454
Teleprinter No.:

3. Further observations, if necessary:

4. A copy of this notification has been sent to:

[X] the receiving Office
[ ] the International Searching Authority
[ ] the International Preliminary Examining Authority
[X] the designated Offices concerned
[ ] the elected Offices concerned
[ ] other:

The International Bureau of WIPO
34, ch min des Colombettes
1211 Geneva 20, Switzerland
Facsimile No.: (41-22) 740.14.35

Authorized officer
S. De Michiel
Telephone No.: (41-22) 338.83.38

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF THE RECORDING OF A CHANGE

(PCT Rule 92bis.1 and Administrative Instructions, Section 422)

To:

GRAVELLE, Micheline
Bereskin & Parr
40th floor
40 King Street West
Toronto, Ontario M5H 3Y2
CANADA

Date of mailing (day/month/year)
14 June 2001 (14.06.01)

Applicant's or agent's file reference
6850-226

IMPORTANT NOTIFICATION

International application No.
PCT/CA00/00430

International filing date (day/month/year)
20 April 2000 (20.04.00)

1. The following indications appeared on record concerning:

☐ the applicant ☐ the inventor ☒ the agent ☐ the common representative

Name and Address

CHARI, Santosh, K.
Orange & Chari, Suite 4900, P.O. Box 190
66 Wellington Street West
Toronto Dominion Bank Tower
Toronto-Dominion Centre
Toronto, Ontario M5K 1H6
Canada

State of Nationality:
State of Residence:
Telephone No.: 416 601-8440
Facsimile No.: 416 601 8454
Teleprinter No.:

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

☒ the person ☒ the name ☒ the address ☐ the nationality ☐ the residence

Name and Address

GRAVELLE, Micheline
Bereskin & Parr
40th floor
40 King Street West
Toronto, Ontario M5H 3Y2
Canada

State of Nationality:
State of Residence:
Telephone No.: 416 957-1682
Facsimile No.: 416 361-1398
Teleprinter No.:

3. Further observations, if necessary:
A new agent has been appointed. Please also note the new agent's file reference.

4. A copy of this notification has been sent to:

☒ the receiving Office
☐ the International Searching Authority
☒ the International Preliminary Examining Authority
☐ the designated Offices concerned
☒ the elected Offices concerned
☒ other: CHARI, Santosh, K.

The International Bureau of WIPO
34, chemin d s Colombettes
1211 Gen va 20, Switzerland

Facsimile No.: (41-22) 740.14.35

Authorized officer
F. Baechler

Telephone No.: (41-22) 338.83.38

ATENT COOPERATION TRI TY

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

Commissioner
US Department of Commerce
United States Patent and Trademark
Office, PCT
2011 South Clark Place Room
CP2/5C24
Arlington, VA 22202
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 04 December 2000 (04.12.00)	
International application No. PCT/CA00/00430	**Applicant's or agent's file reference** 00264-0003
International filing date (day/month/year) 20 April 2000 (20.04.00)	**Priority date** (day/month/year) 23 April 1999 (23.04.99)
Applicant FORSBERG, Cecil, W. et al	

1. The designated Office is hereby notified of its election made:

[X] in the demand filed with the International Preliminary Examining Authority on:

09 November 2000 (09.11.00)

[ ] in a notice effecting later election filed with the International Bureau on:

2. The election [X] was

[ ] was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The Int rnational Bureau of WIPO **34, chemin des Colombett s** **1211 Geneva 20, Switzerland**	Authorized officer Nestor Santesso
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 338.83.38

Form PCT/IB/331 (July 1992)

CA0000430

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
00264-0003		

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/CA00/00430	20/04/2000	23/04/1999

International Patent Classification (IPC) or national classification and IPC
A01K67/027

Applicant
UNIVERSITY OF GUELPH et al.

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 7 sheets, including this cover sheet.

   ☒ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of 7 sheets.

3. This report contains indications relating to the following items:

   - I ☒ Basis of the report
   - II ☐ Priority
   - III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   - IV ☒ Lack of unity of invention
   - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
   - VI ☐ Certain documents cited
   - VII ☐ Certain defects in the international application
   - VIII ☒ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
09/11/2000	27.07.2001

Name and mailing address of the international preliminary examining authority:

European Patent Office
D-80298 Munich
Tel. +49 89 2399 - 0 Tx: 523656 epmu d
Fax: +49 89 2399 - 4465

Authorized officer

SCHEFFZYK, I

Telephone No. +49 89 2399 8602

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/CA00/00430

I. Basis f th report

1. With regard to the **elements** of the international application *(Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17))*:
Description, pages:

1-56	as originally filed	

Claims, No.:

1-57	with telefax of	04/07/2001

Drawings, sheets:

1/58-58/58	as originally filed

2. With regard to the **language,** all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

   These elements were available or furnished to this Authority in the following language: , which is:

   - ☐ the language of a translation furnished for the purposes of the international search (under Rule 23.1(b)).
   - ☐ the language of publication of the international application (under Rule 48.3(b)).
   - ☐ the language of a translation furnished for the purposes of international preliminary examination (under Rule 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

   - ☐ contained in the international application in written form.
   - ☐ filed together with the international application in computer readable form.
   - ☐ furnished subsequently to this Authority in written form.
   - ☐ furnished subsequently to this Authority in computer readable form.
   - ☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
   - ☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

4. The amendments have resulted in the cancellation of:

   - ☐ the description, pages:
   - ☐ the claims, Nos.:

☐ the drawings, sheets:

5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

   (Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.)

6. Additional observations, if necessary:

IV. Lack of unity of invention

1. In response to the invitation to restrict or pay additional fees the applicant has:

   ☐ restricted the claims.

   ☐ paid additional fees.

   ☐ paid additional fees under protest.

   ☐ neither restricted nor paid additional fees.

2. ☐ This Authority found that the requirement of unity of invention is not complied and chose, according to Rule 68.1, not to invite the applicant to restrict or pay additional fees.

3. This Authority considers that the requirement of unity of invention in accordance with Rules 13.1, 13.2 and 13.3 is

   ☐ complied with.

   ☒ not complied with for the following reasons:
   see separate sheet

4. Consequently, the following parts of the international application were the subject of international preliminary examination in establishing this report:

   ☒ all parts.

   ☐ the parts relating to claims Nos. .

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims	4,6,9,11,12,14-16,18-28,30-48
	No:	Claims	1-3,5,7,8,10,13,17,29,49-57
Inventive step (IS)	Yes:	Claims	4,6,9,11,12,14-16,18-28,36-43,48
	No:	Claims	1-3,5,7,8,10,13,17,29-35,44-47,49-57

Industrial applicability (IA)	Yes:	Claims	1-18,29-57
	No:	Claims	19-28

2. Citations and explanations
see separate sheet

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:
see separate sheet

SECTION IV----------

The IPEA is of the opinion that present claims are directed to different problems which are not linked together by a single inventive concept as required by Rule 13.1-13.3 PCT:

The first problem which can be identified is the provision of animals containing/expressing a phytase gene and constructs suitable to prepare such animals which are selected from the group consisting of pigs, goats, sheeps, cows, horses, fish and poultry (claims 13-18,26-28,33,34,36-43, 44-57).

The second problem can be seen in the provision of a method suitable for expressing a protein in the gastrointestinal tract of pigs, goats, sheeps, cows, horses, fish and poultry (claims 19-25.29) and, finally

the third problem which can be identified is the provision of a method suitable for producing a protein via salivary gland secretion in any non-human animal (claims 30-32 and 35)
The only common concept underyling the subject-matter of all claims can be seen in the use of non-human animals to express polypeptides. This concept, however, is not considered suitable to establish unity of presently claimed subject-matters since said concept is already well-known in the prior art. Therefore, the IPEA holds the view that present application does not meet the requirements of Rule 13.1-13.3 PCT.

SECTION V----------

Novelty:

Present claims 1-3, 5,7,8, 10, 13, 17, 53 and 54 do not contain a **technical feature** which would be suitable to render the claimed animal/animal cells clearly distingiushable from any natural occurring ones. Concerning claims 5 and 7 it is noted that the expression "derives from" does not convey an exact meaning but encompasses any PSP or PRP since eventually any sequence may be considered to be derived from another sequence.With respect to claims 13, 17, 29, 53 and 54 which require that the gene encodes phytase it is noted that according to the

information given in the application on page 1, lines 23 and 24 cows express phytase at a low level. Thus, claims 1-3,5,7,8,10,13, 17, 29, 53, 54 do not meet the requirements of Art. 33(2)(3) PCT.
In addition, claims 49-52, 55-57 also lack novelty and inventive step since phytase isolated from E.coli as well as antibodies directed against said enzyme are already taught in WO 99/08539 (1). Crossreactions thereof with the phytase expressed by a nucleic acid sequence according to SEQ.ID.NOS. 3, 5 and 7 cannot be excluded. For the sake of completeness it is noted that even if presently claimed antibodies were novel over those available in the prior art the presence of an inventive step still could not be acknowledged since APPA of E.coli already was available at the filing date of present application. The provision of antibodies directed to a known protein, however, lacks inventive activity since it only involves standard methods.

Inventive step:

Claims 44-47 concern nucleic acid constructs containing a gene encoding phytase and one regulatory sequence for gastrointestinal tract specific expression of said gene. However, taking into account that the preparation of such constructs and the expression thereof in salivary glands is already suggested in WO 97/48812 (2) (see page 18, lines 27-31) the subject-matter of claims 44-47 and claims 30-35 which relate to a process for producing a polypeptide which involves the expression of said polypeptide in salivary glands cannot be considered to be inventive. Concerning claims 30-35 it is noted that it appears that the inventive step establishing feature resides in the unexpected successful expression of a polypeptide in a specific animal (see below). However, claims 30-35 do not specifiy the non-human animals used for expression of the protein. Thus, they do not contain a feature suitable to establish the presence of an inventive step.

According to the Applicant it is not possible to make reliable predictions concerning the successful expression of polypeptides in the salvia or gastrointestinal tract of non-human animals without demonstrating the gene function in the animal, in particular when the non-human animal used to express the polypeptide is larger. The IPEA basically concurs with the Applicant in this position. However, with respect to this the IPEA would like to point out that present

application only exemplifies the expression of APPA in pig! Thus, for obvious reasons present claims cover subject-matter which is evidently not enabling disclosed in present application and hence unworkable (Art. 5 and 6 PCT).

SECTION VII----------

1). Newly-filed set of claims extends beyond the content of the application as originally filed since according to the application as filed the transgene constructs had to be heterologous. However, according to new claims said constructs also can be homologous (Art. 34(2)(b) PCT).

2). In addition, no basis can be found in the application as originally filed for the subject-matter of claim 42 (Art. 34(2)(b) PCT).

3). With respect to the expression "incorporated by reference" Applicant's attention is drawn to Guidelines C-II 4.4 and 4.17 PCT.

SECTION VIII----------

1). It is highly questionable whether all animals specified in claims 1, 13 and 36 actually are suitable to successfully express functional polypeptides (see also above, section V).

2). The language used in claim 48 ("includes") creates ambiguity since it is unclear whether the molecule according to claim 44 contains one of the sequences recited in claim 48 additional or not.

3). Claims 19-28 relate to subject-matter considered by this Authority to be covered by the provisions of Rule 67.1(iv) PCT. Consequently, no opinion will be formulated with respect to the industrial applicability of the subject-matter of these claims (Article 34(4)(a)(i) PCT).

THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:

1. A transgenic non-human animal that carries in the genome of its somatic and/or germ cells a nucleic acid sequence including a heterologous transgene construct, said construct including a trangene encoding a protein, said transgene being operably linked to a first regulatory sequence for salivary gland specific expression of said protein.

2. The animal of claim 1 wherein said first regulatory sequence comprises a saliva protein promoter/enhancer sequence, whereby said animal expresses said protein in its saliva.

3. The animal of claim 1 wherein said animal is a mammal.

4. The animal of claim 3 wherein said animal is chosen from the group comprising pigs, goats, sheep, cows, horses, rabbits, rodents, cats and dogs, and in addition, fish and poultry.

5. The animal of claim 1 wherein said saliva protein promoter/enhancer sequence comprises a parotid secretory protein (PSP) promoter/enhancer, a proline-rich protein (PRP) promoter/enhancer or a salivary amylase promoter/enhancer.

6. The animal of claim 5 wherein said promoter/enhancer is a parotid secretory protein (PSP) promoter/enhancer.

7. The animal of claim 6 wherein said parotid secretory protein (PSP) promoter/enhancer is derived from a mouse.

8. The animal of claim 5 wherein said promoter/enhancer is a proline-rich protein (PRP) promoter/enhancer.

9. The animal of claim 8 wherein said proline-rich protein (PRP) promoter/enhancer is derived from a rat.

10. The animal of claim 1 wherein said transgene is further operably linked to one or more second regulatory sequences including enhancers, transcription regulatory sequences, termination sequences, and polyadenylation sites.

11. The animal of claim 1 wherein said transgene comprises a gene encoding a protein having phytase activity.

12. The animal of claim 1 wherein said transgene encodes a phytase or a homologue thereof.

13. The animal of claim 1 wherein said animal is a pig, said transgene comprising a gene encoding a protein having phytase activity and wherein said first regulatory sequence comprises a parotid secretory protein (PSP) promoter/enhancer or a proline-rich protein (PRP) promoter/enhancer.

14. The animal of claim 1 wherein said transgene construct comprises a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.

15. A transgenic non-human animal that carries in the genome of its somatic and/or germ cells a nucleic acid sequence including a heterologous transgene construct, said construct including a trangene encoding phytase or a homologue thereof.

16. The animal of claim 15 wherein said transgene is operably linked to a first regulatory sequence for salivary gland specific expression of said phytase.

17. The animal of claim 16 wherein said first regulatory sequence comprises a parotid secretory protein (PSP) promoter/enhancer, a proline-rich protein (PRP) promoter/enhancer or a salivary amylase promoter/enhancer.

18. The animal of claim 17 wherein said animal is a mammal.

19. The animal of claim 18 wherein said phytase or a homologue thereof is expressed in saliva or in the gastrointestinal tract of said animal.

20. The animal of claim 15 wherein said transgene construct comprises a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.

21. A method of expressing a protein, the method comprising the steps of:

a) introducing a transgene construct into a non-human animal embryo such that a non-human transgenic animal that develops from said embryo has a genome that comprises said transgene construct, wherein said transgene construct comprises:

i) a transgene encoding said protein, and

ii) at least one regulatory sequence for gastrointestinal tract specific expression of said protein,

b) transferring said embryo to a foster female; and,

c) developing said embryo into said transgenic animal

wherein said transgene is produced in the gastrointestinal tract of said animal.

22. The method of claim 21 wherein said regulatory sequence provides for salivary gland or pancreatic gland specific expression of said protein.

23. The method of claim 21 wherein said regulatory sequence provides for salivary gland specific expression of said protein.

24. The method of claim 23 wherein said salivary gland is a parotid gland, submaxillary gland, or a submandibular gland.

25. The method of claim 23 wherein said transgene is expressed in the saliva of said animal.

26. The method of claim 21 wherein said transgene is heterologous.

27. The method of claim 21 wherein said at least one regulatory sequence comprises a salivary protein promoter/enhancer sequence.

28. The method of claim 21 wherein said protein is a glycoprotein.

29. A transgenic animal adapted for expressing a protein according to the method of claim 21, or a progeny thereof.

30. The method of claim 21 wherein said protein is a phytase or a homologue thereof.

31. The method of claim 21 wherein said transgene construct comprises a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5, or SEQ ID NO:7.

32. A process for producing a protein comprising the steps of:

a) obtaining saliva containing said protein from a non-human transgenic animal, said animal containing within its genome a transgene construct, wherein said transgene construct comprises:

i) a transgene encoding said protein, and

ii) at least one regulatory sequence for salivary gland specific expression of said protein, and

extracting said protein from said saliva.

33. The process of claim 32 wherein said transgene is heterologous.

34. The process of claim 32 wherein said at least one regulatory sequence comprises a salivary protein promoter/enhancer sequence.

35. The process of claim 32 wherein said protein is a glycoprotein.

36. The process of claim 32 wherein said transgene construct comprises a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.

37. The process of claim 32 wherein said protein is a phytase or a homologue thereof.

38. The process of claim 32 wherein said salivary gland is a parotid gland, submaxillary, or a submandibular gland.

39. A method for expressing a phytase or a homologue thereof in a non-human animal, said method comprising:

a) constructing a nucleic acid sequence including a transgene construct comprising:

i) a transgene encoding said phytase or a homologue thereof, and

ii) at least one regulatory sequence for gastrointestinal tract specific expression of said protein, and

b) transfecting the animal with said nucleic acid sequence;

whereby said animal carries within the genome of its somatic and/or germ cells said transgene construct and wherein said animal expresses said phytase or a homologue thereof in its gastrointestinal tract.

40. The method of claim 39 wherein said transgene construct results in salivary gland or pancreatic gland specific expression of said phytase or a homologue thereof.

41. The method of claim 40 wherein said regulatory sequence provides for salivary gland specific expression of said phytase or a homologue thereof.

42. The method of claim 41 wherein said salivary gland is a parotid gland, submaxillary, or a submandibular gland.

43. The method of claim 41 wherein said phytase or a homologue thereof is expressed in the saliva of said mammal.

44. The method of claim 41 wherein said transgene construct comprises a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.

45. The method of claim 39 wherein said nucleic acid sequence is introduced into said animal in the form of a transgene construct.

46. The method of claim 45 wherein said transgene construct is a nucleic acid molecule.

47. The method of claim 46 wherein said plasmid comprises a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, or SEQ ID NO:6.

48. The method of claim 39 wherein said animal is chosen from the group comprising pigs, goats, sheep, cows, horses, rabbits, rodents, cats, dogs, fish and poultry.

49. The method of claim 48 wherein said animal comprises a mouse or a pig.

50. A nucleic acid molecule comprising a nucleic acid sequence including a gene encoding a protein, said gene being operably linked to at least one regulatory sequence for gastrointestinal tract specific expression of said protein.

51. The molecule of claim 50 wherein said at least one regulatory sequence comprises a salivary protein promoter/enhancer sequence, whereby expression of said protein is salivary gland specific.

52. The molecule of claim 51 wherein said salivary protein promoter/enhancer sequence comprises a parotid secretory protein (PSP) promoter/enhancer, a proline-rich protein (PRP) promoter/enhancer, a salivary amylase promoter/enhancer, or a SV40 promoter/enhancer.

53. The molecule of claim 51 wherein said protein comprises a phytase or a homologue thereof.

54. The molecule of claim 53 wherein said molecule is a transgene construct.

55. The molecule of claim 54 wherein said molecule is a nucleic acid molecule.

56. The molecule of claim 55 wherein said molecule comprises a nucleic acid sequence according to SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4, or SEQ ID NO:6.

57. The molecule of claim 53 wherein said molecule includes a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.

58. An antibody specific to a protein expressed by a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7.

59. The antibody of claim 58 wherein said antibody is monoclonal.

60. The antibody of claim 58 wherein said antibody is polyclonal.

61. A hybridoma secreting the antibody of claim 59.

62. A host cell transfected with molecule of claim 50.

63. A host cell transfected with the molecule of claim 56.

64. A host cell transfected with the molecule of claim 57.

65. The host cell of claim 63 wherein said cell is an bacterial cell.

66. The host cell of claim 64 wherein said cell is an animal cell.

67. A diagnostic kit for immunologically detecting a protein expressed by a nucleic acid sequence according to SEQ ID NO:3, SEQ ID NO:5; or SEQ ID NO:7, the kit including an antibody specific to said protein.

68. The kit of claim 67 wherein said antibody is monoclonal.

69. The kit of claim 68 wherein said antibody is polyclonal.

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference	**FOR FURTHER ACTION** see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.	
00264-0003		
International application No.	International filing date *(day/month/year)*	(Earliest) Priority Date *(day/month/year)*
PCT/CA 00/ 00430	20/04/2000	23/04/1999
Applicant		
UNIVERSITY OF GUELPH		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 4 sheets.

[X] It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**

a. With regard to the **language,** the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.

[ ] the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).

b. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international search was carried out on the basis of the sequence listing :

[X] contained in the international application in written form.

[X] filed together with the international application in computer readable form.

[ ] furnished subsequently to this Authority in written form.

[ ] furnished subsequently to this Authority in computer readble form.

[ ] the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

[ ] the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

2. [ ] **Certain claims were found unsearchable** (See Box I).

3. [ ] **Unity of invention is lacking** (see Box II).

4. With regard to the **title,**

[X] the text is approved as submitted by the applicant.

[ ] the text has been established by this Authority to read as follows:

5. With regard to the **abstract,**

[X] the text is approved as submitted by the applicant.

[ ] th text has been established, according to Rul 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from th date of mailing of this international search report, submit comments to this Authority.

6. The figure of the **drawings** to be published with the abstract is Figure No. _____

[ ] as suggested by th applicant.

[ ] because the applicant failed to suggest a figure.

[ ] because this figure better characterizes th inv ntion.

[X] None of th figures.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/CA 00/00430

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 A01K67/027 C12N9/16

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 A01K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
STRAND, EPO-Internal, WPI Data, PAJ, MEDLINE, BIOSIS, EMBASE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 97 48812 A (UNITED KINGDOM GOVERNMENT ;YANKE LINDSEY JAY (CA); CHENG KUO JOAN) 24 December 1997 (1997-12-24) page 18, line 27 - line 31 ---	1-69
X	WO 99 17610 A (UNIV CALIFORNIA) 15 April 1999 (1999-04-15)	1-5,10, 50-52,62
Y	page 7, line 12 - line 26 page 10, line 21 -page 11, line 22 --- -/--	6-9

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 September 2000	13/10/2000

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Sprinks, M

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
[redacted]T/CA 00/00430

C.(Continuation) DOCUMENTS CONSI[redacted]D TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	SELINGER, L. B. (1) ET AL: "The rumen: A unique source of enzymes for enhancing livestock production." ANAEROBE, (1996) VOL. 2, NO. 5, PP. 263-284. , XP000940728	21,22, 26,28, 29,50
Y	abstract page 267, column 2 page 276, column 1 -page 277, column 1	15,30, 39,40, 45,46, 48,49
X	ZHANG J X ET AL: "Expression of a bacterial endo (1-4)-beta-glucanase gene in mammalian cells and post translational modification of the gene product" BIOCHIMICA ET BIOPHYSICA ACTA-MOLECULAR CELL RESEARCH, (27 JUN 1997) VOL. 1357, NO. 2, PP. 215-224. PUBLISHER: ELSEVIER SCIENCE BV, PO BOX 211, 1000 AE AMSTERDAM, NETHERLANDS. ISSN: 0167-4889., XP000940605 abstract	50,62
Y	WO 99 08539 A (DIVERSA CORP) 25 February 1999 (1999-02-25) Discloses a phytase - relevant for present SEQ ID NO:7 figure 1	11-20, 30,31, 36,37, 39-49, 53-61, 63-69
Y	DATABASE EMBL 'Online! ID: AF062078, 18 May 1998 (1998-05-18) DENNIS ET AL.: "Cloning vector p34S-Cm" XP002148835 Relevant for SEQ ID NO:4 abstract	56,63,65
Y	DATABASE EMBL 'Online! ID: MMPSPG, 3 November 1992 (1992-11-03) MIKKELSEN: "M. musculus Psp gene for parotid secretory protein" XP002148836 Relevant for SEQ ID NO:1,7 abstract	1-7, 10-14, 16-69

-/--

1

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y	DATABASE EMBL 'Online! ID: BLCAT3DNA, 26 February 1992 (1992-02-26) LUCKOW: "Plasmid pBLCAT3 gene for beta-lactamase and CAT gene for chloramphenicol acetyltransferase" XP002148837 Relevant for SEQ ID NO:2,6 abstract --- DATABASE EMBL 'Online! ID: MMU73190, 4 November 1996 (1996-11-04) LAURSEN ET AL.: "Mus musculus positive acting regulatory region of the parotid secretory protein (PSP) gene" XP002148838 Relevant for SEQ ID NO:7 abstract --- DATABASE EMBL 'Online! ID: CVCATBLA, 24 January 1992 (1992-01-24) BOSHART ET AL.: "Cloning vector pBLCAT6 encoding chloramphenicol acetyltransferase (CAT) and beta-lactamase (bla) genes" XP002148839 Relevant for SEQ ID NO:3,5 abstract --- DATABASE EMBL 'Online! ID: RNRP15, 12 May 1991 (1991-05-12) LIN ET AL.: "Rat salivary proline-rich protein (RP15) gene" XP002148840 Relevant for SEQ ID NO:3,5 abstract --- DATABASE EMBL 'Online! ID: ECAPPAA, 5 February 1991 (1991-02-05) DASSA ET AL.: "E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene" XP002148841 Encodes a phytase - relevant for present SEQ ID NO:3,5,7 abstract -----	47,56, 63,65 1-7, 10-14, 16-69 14,20, 31,36, 44,47, 57-61, 64-69 1-5, 8-14, 16-69 11-20, 30,31, 36,37, 39-49, 53-61, 63-69

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
[redacted]/CA 00/00430

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9748812 A	24-12-1997	US 5939303 A	17-08-1999
		US 5985605 A	16-11-1999
		AU 716845 B	09-03-2000
		AU 3021697 A	07-01-1998
		BR 9709793 A	11-01-2000
		CA 2257101 A	24-12-1997
		EP 0904385 A	31-03-1999
		NO 985804 A	10-02-1999
		PL 330532 A	24-05-1999
WO 9917610 A	15-04-1999	US 6004944 A	21-12-1999
		AU 9668998 A	27-04-1999
		EP 1030557 A	30-08-2000
WO 9908539 A	25-02-1999	US 5876997 A	02-03-1999
		AU 8904498 A	08-03-1999
		EP 1003379 A	31-05-2000
		US 6110719 A	29-08-2000

SEQUENCE LISTING

<110> University of Guelph
Forsberg, Cecil W.
Golovan, Sergeui
Phillips, John P.

<120> Transgenic Animals Expressing Salivary Proteins

<130> U Guelph - 03

<140>
<141>

<150> 60/130,508
<151> 1999-04-23

<160> 7

<170> PatentIn Ver. 2.1

<210> 1
<211> 20623
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Lama2/APPA plasmid

<400> 1
tcgagagtat ctttgtcagc tgtgcctcca acaaaggggt actgttgccc acatagaaag 60
atctaaacta attaattaat ccctcacccg caaatctttc agtcactaag ttagcacgat 120
tgttgaacaa gttctccaaa ggagagatac agatgagtgc gtatagggtg gacctggctg 180
ctgaggagac acctgcatct gactaagaag agccacggtg ttagttgaat ggtgtggagt 240
agggtggttc tgtgggacag tagaaaatcg agaggcatgt gccgtttagt gaactgatgg 300
aagctacccc aaacgacaga gattgtcagt caggccaatc cgtttcgagt ttgatgggca 360
gccggacagt gagacagaca cacctactca gttggaggaa ggatgagaac aatggccagc 420
agggattgag agaccctgac aggcgcaagg ccctaacaca cacacctacc acctcacttg 480
acaaagctgc caaagaccaa agacttgttc tccattagaa atgacagctg gcttgacccg 540
acagcataat aagcagagtg tactctgatt ggagaacttt aatgtgtttc attcagtatt 600
ataaaaggac agtattacag attttgttgt acactgctgt tacatgtggg gcagtgtgtc 660
tttaagtagg gtaaagtact ctttaaaaat gggtcctaga tatttttttcc tttaactcaa 720
gtctcttact gtttaaatga tttttatttt gtttaatatg gaggaaaaag aagcgtaaat 780
ggacaatata tatttagaga aagatggtta gctgtcagaa aaatatgcaa atcaaaatca 840
caccaagact gcagcacacc cctgtcagat ggctgtgatc aagaaaataa atgacaatga 900
gtggtggtga agatgtacta aagggaaaca cacacacaca cacacacaca cacacacaca 960
cacactggag caaccactgt ggaaatcagt atgaatggtc ctcaaaaacc tgaagataga 1020

gcggggcgtg gtggcataca cttttattcc cagcactggg gaggcagagg caggtggatc 1080
tctgagttcc aggccagcct ggtctatagc acaggttcta ggacagccag ggctacacag 1140
aaaaaccctg ccttgattaa accaaaccaa accaaaccaa accaaaccaa accaaaccaa 1200
accaaaccaa accaaaccag accaaaccaa aacactgaag atagaacttc agtattccat 1260
tcctagatat atacccaatg gagactaagt cagcaagaca cctgcacagc catgttcact 1320
actacactgt tcaccacagc caggctgtgg aaccagcctg agtgtccatg ataaatgaat 1380
ggataggtaa ctttcaaggt aaatggactc tgctgtgtac atgcctcaca ttctgtttat 1440
tcatttttct ttatgaggtg tccattcagg agtcacatgg tagttctatt ttcagtcttc 1500
tgaagatact acactggtcc ccacagttta cacttttatc agcagtgaat aagggttcct 1560
ctatccttac catcatttgt tgtaattttt cttgatgacc ctctttctga cagggatagg 1620
atgtaatatc agtgtgagga agtacaactt gttttctaag tatttattgg ccccttgcat 1680
ttcttctttt gaaaactgtc ggttcctgac atctgctcag gtattcattg gatgttgttt 1740
ctttggtgtt tgagttctta tgaattctag atgttaaatc cctgcctgtg gttctctccc 1800
attctgtagg ctgcctcctc accctggcaa ttgttgtcct tgttttgcag aaacttttga 1860
cttcatggaa tctcatttgt cagttttccc tcctctgcta tagcctgagc taatgcactg 1920
gtttttacag agccctggtc tatgccttta tcctcctctg gcagcttcgg agtttcattt 1980
cttacattta gatctttgat ccactttgaa caagttttgg agcagggtga gagatacgaa 2040
tctagttcca ttcttccata tgtgatccta gtttacatag catcgttggt tgaagaggtt 2100
ttattttatt tttaaataat gtgtcataaa aaacgaggtg gttgtagcag tgtggatttg 2160
tttctttgtc ctttgatcta caggtcttgt tttgtgtcag tctcatgatg ttttattgct 2220
atggctctgt catacagtct gaggtcaggt attgtgatat accttcagta ttgctccctc 2280
agactcaggt ttgctttggc caggagtcat cttactcagt gctcttagag ctcccccagc 2340
atgtagctgc tactattctt agttgataaa tcaggaaact ggggctcaga gagattaact 2400
gtcttgaact acttctgggg aggtgaaacg tggagacact aaactgtgtt taccctgtac 2460
tgctccagta gctgtcgggt gctgggctac agcaaagcac ctatactata tattactcag 2520
gaggtggaaa aactcagcct cccttggggt tcccaagctc ccaggtgtcc agtcactgct 2580
ggaaacctca tggagtctga aaggaagggt tgagggtaca tggggcagcg atgaggagcc 2640
tggggctggg atctcccaaa cacctggata tccagatgcc actgggtcag ggggagttgg 2700
gaacagagtt gggatgtcca tggacctgtg acaaggccag ggccaggggg aggataactc 2760
tggctttact aatttgcgaa agtccttagc ttagcagcag ttgtctggga gcacagaggg 2820
gccttctgta agaggctcag gcagtgccgc tctgtaggcg aaggtcttct ccatgttccc 2880
catggtggtt cttgatgaaa gagacagtcc ttggctccaa actggtttat tgattgttca 2940
ttgtggaaaa tgggtgcaca ccaccttctc agggtggacc agagatcaaa taccttttgc 3000
agggaggaat atctgggaag ggacgcttac tggctaaacc ctcagggcct ctagatacat 3060
cattagcatg gagaactctg ttctgggcta catgaccaca ggccacattt ccacaagcca 3120
catgtgggaa gtgtggcaca tgttctaggc caggaatctg gtagggagcg tggagccacc 3180
taccatccca ggtgggtgcc tgggtgccag ggaccctgaa cccgctcaac cttaccaagt 3240
ttcctggcag ggtccactgt cctacacaga agctggagga ggtgtgaggg ttgtgtcttt 3300
gtggaatgtc ccatgctgct tggggctcag tttctccacc tgtacctcat tggtttgggt 3360
ataaaaagtg gggatacttt attattctct gactcggtcc tgaggaaaaa gcatcgtggc 3420
agtccaggaa ccacaccctg aggttcctgc actgaaggga ctccctaagt ctctggagtc 3480
tctccccttc acagagctgc caaagtctag gttcttttga ggataacaga gccatgcttg 3540
gtaagcagac aacagcattt gtttactcaa ccttcttttg tcagctccct cttcataaac 3600
aagttgagac accatgctgg cttgaggaag acttctaaag ccagacaact gtgcaaggaa 3660
gaagaagaag gggcaagtgg agttagcctg gatgtagccc tcaaagtctc cagagaccag 3720
ccatgaaggc tcaagtggag ggcaagacct gcagcagcca agcatctggc aggagaggat 3780
cctgggaacc cctctaccat gacacacatt cttcctgcag gtcacactta ataggccatt 3840
tcttatttgg atctatcatg gtgttctgtg cgagattaat gaggtgttat gctgcgaaca 3900

gaaagttata taaaaacaag tccccccccc ttgtcactgc tgctaagaat gtagcagaaa 3960
ttgtctcaag tgtctctcta atcagaaaca ataaaggtct ccttggattc aagccctcca 4020
gtttcctcct tccttgctga gccttggaca cccatacaaa cctcctggat gctacagctc 4080
tgggcagaga ctccaaggtg gggagagact gatggtacaa aagcaaaata cttgtttggg 4140
ggtacaccca ctcctctgcc tgtgtggttc ctgcagtcag tcctgcagac aggccctcag 4200
tgggtcttcc atgggcaaca cgcagaggga ggcaatggat gggaataccc acaccctggt 4260
tagtttaccc cggccatgct ctctgctctt catccctcct ctgccctctg ccacggcttt 4320
ctctgcagga atcatatctt catattggcc cacaggtgtt ctcctcaccc tagctatgat 4380
gtttacttta gagtgacctt agcagggctg gtgggaatga gttctagaag gctcacggag 4440
atgctaggga agaaacgtct tctaactact gaggttacta agttcctggt ggttgtctct 4500
gcctttccct tgttaaagtc accttgaagt tagtgcagaa gaaatcagag cccagtcaca 4560
gagtaaatat ggtcctgaag atttcctttg agtgcccaga atccatgaca tttcaagagc 4620
cctctttgta ccttaagtca tttggggttg tatcttctgc ttgatgtatg tgtgtgtgtt 4680
tatcaaagag tgagatggtt acataagagg tgctctaaag gacagagagg atttgcaatt 4740
gtggcatgtg acatcctcag gccttgctct ggtgccagga ggaactgatg cagaaaagag 4800
taagaggtca tttcctggag gctgtcacta tagaggagat cttacagtgc attccctcct 4860
ccaggccctg cctgaggata gacatgtgct gactgcaact gaaacagagg cttgggatgg 4920
agagttaggt tcacagaagg gagggtggga gatggatgct tgctgggttc tgggtctcat 4980
caccagctcc tgaccacccg gtcagcccat gtgcttattc catagctttc ttttgctatg 5040
tttactcagt gtggtgtttg ttgggaccca gcagaagcca gtcccaggct gacagctgtg 5100
gatacacagg gcagcatgag ggtcctcagc ctgaagcagt caggctggca gaagagaaag 5160
accagcacac attccttcaa ccaactatgt cttgaaaaac aaacatatta tatcacatat 5220
attgcattta tgagacagct aaaatgtact cgggtagcat gactccaggt ggggatatct 5280
gcaagtgcca tgagtggcag agggacagcc aatgtgaggc aagaaggaat tctggctcaa 5340
cacagcttag ctccctggtg ttggttcaaa ctttgagagt ttgaccacaa gcactttatt 5400
tttgacatat ttaaacagag cacaactttg ggaaaaagtt ttcttatgaa aattatcaca 5460
ataaagctta aggcatgact acattaaaat gcctttgcaa agtatatgtg ccctcttcca 5520
caagaatggt tctattgact gagaaataat gttcaggata aagatccagg aagaaaagat 5580
cagggataag taaaatacta aactcttttg caaagtacat agaccctctt tcataacaat 5640
gggttctatt gactgacaag cactgctcag gagttgggaa agagtctagc ataagcacga 5700
tagcctggag actctagtga ggtctagtct tacagacagc aaaaatcacc aggttacaaa 5760
ctacattcat ttccagtttt ctgatcaggc acaggtatga atcccttctg ttgaagagaa 5820
aagtccatgt gtttaaaata tctggtttct ccagtgctat tagcgagaag acttgagccc 5880
tatacaactc ccacctggag tgacatcctg tcttcatggt atattacata cctagacacg 5940
ctcatctcac agacttagga ctttgtcttc tgatctccat ttctgatccc acttccacct 6000
ttgccttgat agtgtcattt tcttcactgc cttggtgaca accatgttat cctctgtgta 6060
tttgagtgtt accattttca gattttacct gtatgcaaga tcacacagtc tttgtctttc 6120
tgtctggatg catgctaatc tctacacaac aacccttccc cgtcactcag atcttcctcc 6180
attaacacat acatggtgct gaagaggcta gggagcttcc cttcagtggg gagctagctg 6240
gctattgggc ctttttgact gtccaggaag gcccccaatt gctgagacaa gaacttagat 6300
tcttcattat tgactctaac tcatgtatca agcagaagct aatgaatagt tatcaacagg 6360
atcagaggtt ccagtgtaag acactttgac atgaaagaac ggaggaagga cagatggatg 6420
cataaaagca ggaccactgc cccaggaagg tcctggaaac tgatgcaggg caaaggacag 6480
gttataaacc aaatcttagg gagtcaggaa gagcacagag gagctcaacc aactgaccac 6540
tgcttagggg ctaccaaccc aatcctccct gtgggaacag ctaagctatc agccaagggt 6600
aataaacagg caggacctgt ggatgacatg gagagcatag ggaccctggg tccagccttt 6660
agcacctgca ctctcaggat actccaccat tgtgtcttag agagcctagg gatactgggt 6720
ccagcctttg gtaccttcac tctcagggta ccccatcact gtgtcttgga gagcctaggc 6780

accctgggtc cagccttcag tacctgcgct ctcaggacac cccaccattg tctcttgccc 6840
cgtctcttct tcctcttcct ccctttcatt gtctcttctc tgtttctttc ttgactctcc 6900
tttcccctca caccctcact ctagttctcc ccttccctct ctgcatcacc ctattctctc 6960
tgtggtccct ccactttcct ttatctctca tgcttctctc ctccctcaaa tacttgtcac 7020
ccactatact tcaggggcca gctctagtga caaagctgtt aatagcaaga ctctcagatc 7080
tccaacggct cagaggagcc agacccacca agaactctct ccaggtccaa tttcaggttc 7140
cttcgaaagc tttcagcaaa tgctcaggga acatgccact aacaagaaga tgcaaattcc 7200
agttgagagt gggaaaggcc cttgcgtagg tcccatcttc caggccaagg tcagaggggc 7260
tctgtgtaat ccggattgac agggctcaga acaatgtttt gtttttaagg tttatttatt 7320
ttaggtgtta gtgtctttgc ttgcatgacc ttatgtgcat catgtgtgtg caggttcctg 7380
atgacagtag aggagggctt tgaatccctg gggataggaa gttacaggaa attataagct 7440
gctttgtggg tcttctagct ttcccaacag aagtgaatgc tcttcaccac tgagccatct 7500
ctctaggccc aagagacatt gctttatgga tataattgtg tgtgtgtgtc aacattgagg 7560
aaagggaaat aaaaaaaaaa cttcagccgc taaggttgta cagtttcact aattgctact 7620
tttagttgtg ataaaatggc aggtgcttca acatttatat atacaaaaac ttccctgctg 7680
gtggttcaac tgtgagaact ggggtaagtg ggtgagttct ctttttctgt ctctgtctct 7740
gtctctctcc ttccattctt tcttaaagga aataaacatt gcagctgggt tatagctcat 7800
caatatggaa gttacagaag tgaaaaaagg cattgccttg gtgggtggtg ttaccagctg 7860
atttttggtt gtcctgcaag gaggtctggg gactggctgc tctgtctctg tctgtatgag 7920
tgagggaagt ctggggagca gattccctaa ccttcagcct ggcctggttc ctgagtgaac 7980
ccagcctctc tggtcctagt agctttttcc aaacaggaat ctgagtggtg acagggaaca 8040
agtaccagcc cattgcttaa gtgccagggt tagtgagggc aggaagctgc catagctggg 8100
attagtagtt gtattggatg taggaagtcc tatcctggga cagctaatcc ttaatgcttc 8160
actggagatt ttcaatgaga aatttatccc acggcccata tggccccatc cttttgtctc 8220
caacagccaa gtattttcca ttagaggaga cttcctgtac acttgatgga tgctcattcc 8280
aaggtgactt ggggcagtca gtacagactt gggatgacct ctgacagcct aacctctccc 8340
caacaagggc cctctatgtt tgctatgtaa tgtaatgtca gacattgtca ggagtgtccg 8400
cagcacagcc tgcccagtgt gagggctctc ataggtttcc cactgtctta tctacacagg 8460
gataacgagg aggtaagctg cagttcccag tctcacttca cagaggaaga gataacccca 8520
tcccaggtca tgtagccagc agtggaaaga atgaggattt gaactcaggt cttccaagtc 8580
ccattgatag catctcctca caagtccctt gccaccctca cgatgcctta gacacttgcc 8640
tgccctttat actaaggaga tgcaggtaca aggggtttac ccatgtagca gctgaggcag 8700
ctggggatag ataccagcag caggcctgat gtcaccactc taactccagc atccccagtc 8760
tgtgttcctg gagtgtgaaa atccctactt aacaagattg tgcaacagtc cttggctctg 8820
tgacccatag ctggaaacag gattctcatt gatttgtgga acatggtggc agccagccaa 8880
aaagagggtc tgcatacaga agacacgtgt ggcaaggcca cagcagactc tgactacctt 8940
agcttacaga attacaaggt cataatgtcc tctgctttgg tcacctcatg ttaaggacag 9000
gccctaatga agatggggca gaagactgaa ggaatggcca accaataact ggcccaactt 9060
gagacccatc ctacaggcaa gcatcaattc ctgacactac taatgatact ctgttatgct 9120
tgcagacaga agcctagcat aactatcctc cgagaggtcc acccagcaac tgactgaaac 9180
agaaaaagat atccacaggc aaacagtgga tggaggtcag ggactattat gggagagctg 9240
tgggaaggat taaaaaccct gaaggggata ggaaccccac aggaagacca acagagtcaa 9300
ctaagagacc tgtgggagct ctcagagact gagccaccaa ccaaagagca tacacaggcc 9360
ggtccgaggc acctggcacg tgtgaagcag acatgcagct cagtctccat gtaggtcctc 9420
caataagcgg tagcctgact gcagtatcca atccccaaca gggctgcata gtctggcctc 9480
agtgggggag gatgccccta atcctgcaga gacttgatga gtggagagct atccaggggg 9540
aacccaccct ctctgagaag ggaatgggga tgggggaggg actctgtgaa gaggggacaa 9600
ggacaaacaa gaacctcaaa taggtcaggc cctaaaggct tgctaagtag cagtggccca 9660

gctctgtcct gttcctcagc ccaaggctca gctcccacct gtttctgtgt ttttctggct 9720
tttcatgggc ctaggacttg gtgaccagtt caaacaatgg ggcctgtgga agacacaata 9780
tacaagacta gggacattcc tgttctgctg actatccata gcctgatgta ggtggaagga 9840
cccaatcact ggatttctac ccttgcacaa ccttgacagc tgagggcctc tcagaaacct 9900
atttcttcca ctgaaaaatg agactctcaa atgaacgtcg tgacaatcat caggcttatt 9960
aaagaggtgt atctaacctg aatggcaagc agacagcagg caaatgtctg tatcaacctc 10020
taggaaggac aagaactgct cactgctgcc ccccaggagg ccatttgctg aaacagctgc 10080
tctcctgctg gtgcacaggc cctgccttct cattgcagcc acagcccctt cctgtctgaa 10140
cctcctgtca ggtcactggg aaacagatca agatggaaca ggacagctcc tgatggtaaa 10200
taaaaaacag tggtcatggc tattcatagg ggtttatgct tcttcagtcc acactgtgaa 10260
gagctgtggg catgaaccac agtgttcgag gtagagttgg ggttctgaaa ttcacagtgg 10320
ggtgagctca gtaaatgtga gctggaggtc actcgtgaga cacacagtcc tgctgcttct 10380
gttcccaata tcctgaggag acgacacatc tactttgttc agaggccaca gtctagttga 10440
cctgagagtt accagtttct tatttgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 10500
tgttgttcgt gtgtgagtgc aggtgcacat atgatagcgt acacgttgag gtcagaggat 10560
aactatcagg cgttgtcccc tcctactttt cctcggactc tggagaacaa acatgggtcc 10620
ttattccagg ggagcaagtc gctgttggct gacacatctt gctcacatac attttaccta 10680
gacaatggag cctccatcag agtattactt tagctcctca ccgatggcaa tgcaccacct 10740
ctctacccac ataggagttg ggtctccaca cacccccaca ccccccttcac caaaacgttt 10800
tcagttactt tatctggtaa agttcatcag agaatgaagc cagtattaag aacatggaat 10860
catttgggaa cctggatcta gcaatacccc accctagatg gagttgctga gttttcacct 10920
cagattataa ttccccccta gcttctatgg tttattctga aaccagggga actcgattcc 10980
tccctttgga ccacagacat cctggcttgt gaattcacat gtcatctact gctaatccat 11040
tggtagtatg tggctcacag agacacacta cagtcatggc caatgtcaag gtaggacaga 11100
tgtgaatcat tcccccagtc ctgctgtttt catgactaac cctcctcagc acagtgacca 11160
tgaacctact tttcccctcc ttttatttttt agaattgctg gaattttcta ttttgagaaa 11220
taatagcctt gggcagcatt aaacaaaatc atctagaaag ctggtttaaa atacagatgg 11280
ttgagtcagt gaaagagtga ggaatgtcat tattggcccc tcacagaggc tggctcactc 11340
cagcagaggt ggttgaagct cttggacacg ggtcaggtgc ataggaaagg tngtctggga 11400
cactgagaac cacaattgaa caaacagaac tgttggcttt ttttttttta aatgagttct 11460
caaaaaatga ctggctagct taggcaaata cttcgagcca acccaacaga acattcttcc 11520
attgattcat tctggatctt ctttctagac aatactgaac tgacccctg ttggcagtct 11580
caagtttgac aacatagggc tttgaacttg gcacaaggtc catcactgtc acccaagcat 11640
cctgggtgac ctttgggttg gaatatcttg gctaacctta gatattttct ttggagtatc 11700
tttagaacat ccaggaaata gggcttgatt ctcatcctgg gaccacaata taagtcaccc 11760
tagaatccca ggagatcgtg cagagaaaca aggatctctc tcgtgtgcat ccttcttcaa 11820
agcagtgagt agtgactcca ctaaactgag ttcccatctg agagtccaca ggaggctttg 11880
gggcaagaag cagagggaag gcactgtttg tgttggtaaa gttttgactc taacaaattt 11940
gaagacatag atgacattgt gtcagactaa caacaaccta gactcatgtg ggttctgttt 12000
agggatcaga ttttattcat caatgacttg tcttagtgta tagagaaagg cttcctactg 12060
gagtgtaggc tcaataatga cagaagagat agctatttcc cctagggact gtgctgctcc 12120
aagtttggtg gagaaaggca gtggggaacc tagatgtgct ctctggggag ggggtctgaa 12180
gctggcttca tagaaggtgt gaagttttgc tgaaacatct aaacagaatt atagcttagg 12240
aaagtgagca ggcaaggcag ggaatgtgtt gcatatgtat atgtacatga atatattatg 12300
ttatagatac acacacattt gaacctcatt tgcagatgac agaaaatagg ttattttgcc 12360
tctcttaact gctaagcaca atgacttcca gttccatcca tttcctgaaa tgccacaatt 12420
tcatttttca ttgtggctga ataaaattcc attgcagact gggccctact tcatccactc 12480
ctgagggcag gcatatcccc tggctccatt tcttacctat tgtgaagaga agtgcaactg 12540

tcttgttgaa aggcaagcgt gagagaggca ggcactaatt gtgggttttt gtttcttctt 12600
cctgctatga ctctccattt gtcagaacca aagatcgata aaagccgcca ccatgaaagc 12660
catcttaatc ccatttttat ctcttctgat tccgttaacc ccgcaatctg cattcgctca 12720
gagtgagccg gagctgaagc tggaaagtgt ggtgattgtc agtcgtcatg gtgtgcgtgc 12780
tccaaccaag gccacgcaac tgatgcagga tgtcacccca gacgcatggc caacctggcc 12840
ggtaaaactg ggttggctga caccgcgcgg tggtgagcta atcgcctatc tcggacatta 12900
ccaacgccag cgtctggtag ccgacggatt gctggcgaaa aagggctgcc cgcagtctgg 12960
tcaggtcgcg attattgctg atgtcgacga gcgtacccgt aaaacaggcg aagccttcgc 13020
cgccgggctg gcacctgact gtgcaataac cgtacatacc caggcagata cgtccagtcc 13080
cgatccgtta tttaatcctc taaaaactgg cgtttgccaa ctggataacg cgaacgtgac 13140
tgacgcgatc ctcagcaggg caggagggtc aattgctgac tttaccgggc atcggcaaac 13200
ggcgtttcgc gaactggaac gggtgcttaa ttttccgcaa tcaaacttgt gccttaaacg 13260
tgagaaacag gacgaaagct gttcattaac gcaggcatta ccatcggaac tcaaggtgag 13320
cgccgacaat gtctcattaa ccggtgcggt aagcctcgca tcaatgctga cggagatatt 13380
tctcctgcaa caagcacagg gaatgccgga gccggggtgg ggaaggatca ccgattcaca 13440
ccagtggaac accttgctaa gtttgcataa cgcgcaattt tatttgctac aacgcacgcc 13500
agaggttgcc cgcagccgcg ccaccccgtt attagatttg atcaagacag cgttgacgcc 13560
ccatccaccg caaaaacagg cgtatggtgt gacattaccc acttcagtgc tgtttatcgc 13620
cggacacgat actaatctgg caaatctcgg cggcgcactg gagctcaact ggacgcttcc 13680
cggtcagccg gataacacgc cgccaggtgg tgaactggtg tttgaacgct ggcgtcggct 13740
aagcgataac agccagtgga ttcaggtttc gctggtcttc cagactttac agcagatgcg 13800
tgataaaacg ccgctgtcat taaatacgcc gcccggagag gtgaaactga ccctggcagg 13860
atgtgaagag cgaaatgcgc agggcatgtg ttcgttggca ggttttacgc aaatcgtgaa 13920
tgaagcacgc atacccgctt gcagtttgta aggtacccgg ggatcacaac ttgccctctg 13980
aagaggaaga acagaaggat gccacaactc tcctgctggc tactctccag tggtttcatc 14040
ttacttctga tggcatttcc ctctagaaag tgctactatc atccacacat ttctacctga 14100
gaccacccaa aggaccctcc caaattctct tcctctctga gtagtctcca cacctgttac 14160
caccatccca gaattaaaat cctaactgca ctctggcgtg tgacttgcct cagtccttgc 14220
aataagagtt gttggcagtg ccaggcgtgg tggcgcacgc ctttaattcc agcacttggg 14280
aggcagaggc aggcggattt ctgagttcga ggccagcctg gtctacagag tgagttccag 14340
gacagccagg gctatacaga gaaaccctgt gtcgaaaaac caaaaaaaaa aaaaaaagtt 14400
gttggcagag tgtgggttat ataccaggtg gagatttcaa atgagtggct gaagctgtag 14460
ccagaaggaa cttagaggat agctcataac ttaaaaagaa atgtagagag tagcagaaac 14520
attgagagag tgggcacaca gccactgtgt gaatgtggca gaacacaatc cagccagcta 14580
tacatgcata agtgtatatt ggcgccatcc tgactgatga gacacaggaa aacagataga 14640
cggggttagg tggccatggc ctttcctgcc tgcctcttcc taagggtcat ctcaagacct 14700
tatgctctct taactcttcc attgctactt agcttctaga tatcacctcc agattagtct 14760
ccttgggtac atcagtgatc ctggtgatat ccagggcttc ctgattccat ctttgtcata 14820
gaggctgcaa ctaaagaggt cttcttaata cttcacaccc tgatgccaaa aggaagacac 14880
agaagttcac agaggtgaag tgattcatgt aggacataca gtgagcaagc atcagggtcc 14940
ggattatctg actctactct aacttttatg taaatgtgct ttatgccatt aacactgtca 15000
ttcctgtgct tcagctctgg gagactccca agcactctta ggcacaagcc acaattaagg 15060
gactctgaca ctctgcattg attaattagc atggtggtct ctatgtttcc agattcatga 15120
ttgtttcact ttccatatag gctatgaagg gtgtgaggaa atttttgggg gacagaattg 15180
gaggcaatcc acctctctca ggaagcctct atctggaaaa gcttacaact cagggacagt 15240
aactgtaggc ccagtccttg gtgtccaaaa tgggttttat ggtttgaatc tgcaaagcct 15300
tccatgtgct caaaggtttg aacatggagc ctcctcctgg taacactgta ttggaggctt 15360
ttgagactgg atgctctttg gtcccatgtt ttgctacatc atctgtcaag atatgaccca 15420

ggcatgctac cagctaccac agactatgcc tctccagctt tcatgttctc cccaccatga 15480
tagacttgta tctcctaaaa atggaatcaa agcaaacttt tcctgcatta agtttttttt 15540
tttctgttaa gtgtttggtc acagggacaa gaaaacactc aatacagata attagtacca 15600
gagttgaggt tcattgctct agcaagttgg atcaaatttt tagggctttg gaactgattt 15660
ataagagaca tgtagaagag tctgaagctg tgggctacag aagtgtcacc agtttttaag 15720
aatagtttaa tacaccatgg gaattgtgaa aatcagaatg ctcacacaaa ggcagacagg 15780
aaaacgtgag catgtggcgt gtgagagggc ataagaagga acctaggggg aaatgagcta 15840
gaagccattc ggctacgtta gggaacgtgt gtggctgtgc ttggcccatg ccctggcaat 15900
ctgaatgagg ccaaatttta aaggagtgga ctaactcgat tgtcagagaa aatatcaaga 15960
cagaccacca ctcaggctat gccgtgtttg tgaccgacca gctactctta gccagctcta 16020
ttgtgaaatt ccagagcaat tatcagagca tgaagataca tacagtttag tgaagtaagg 16080
ggtgtgggtc cctaagtgga tggtgcataa atctatgtag gtgatgccta agtgacactt 16140
gataatccaa aatatcagca atgtggaatg tcttccaagg agacctgtag acacacattt 16200
tagaactttg ctcatggctg taataaatag ctagctagaa atcatttcct gaagaggtta 16260
gtctgagtta cggttccagg gcaaacattc agtgatggca aggaaggcat tgcagtcagg 16320
agccaaaggt cagctggtca cattgcatca agagtagaga gtcagagtgt gagtagaaag 16380
aggatacagg ttataaaacc tcactgtcca ctctcagcaa tccatttttct cctaaaaggc 16440
tttaccttct aaagatttta gtcttcaaaa ccagtaccag tagcctggga acaaaagttg 16500
aaacaaatga gcctttgtgg ggcatttcac acttaaaaca gggcatcacc taggaggagc 16560
cctgtgtgca gtaggaagtg tggcctctgt gtcaggaatg ctcaggctaa taaggggtcc 16620
tctatctgag ggaccctatg aagattcaac aagtagttgt gagaattccc tgtaaatgga 16680
tgctaccaat ttgacatttg tagacctgct attgtgtgct tctttattgg gctctcccat 16740
ctcccaactt tccaacccat attccacatt aatcccttcc accaccatgc aacactaggt 16800
aggagagaag gaaggttaga agagaaagtg ggtatagatc tatttagact acttcctgct 16860
gattaggggc aagtccaatc gtcattgtca ggatacctcc aaccagcaac cagcaaacca 16920
gcaaatcaga aacagcaaaa gcagccaaca aggcagcact aaccagcagg attggggtcg 16980
gtagcgtggg agcagtcact actggtcttc tcatggcttt ggcattaata ctctctcaag 17040
aaattccgta atttttccc caccacctga aattccgtaa ttttaaatgc aaactatcta 17100
cagctggcaa aaatcacatc tctcctagag cacaagacaa atcatagtta ctggctattt 17160
gcaatctgaa gcatctcaat atcccacacc tgggattaaa acaaaaacat attcacatca 17220
cataactgtt tttttttcc aatttttat taggtatttt ctttatttac atttcaaatg 17280
ctatcccgaa agtcccctat accctcccac ctccctgctc ccctacacac ccactcccac 17340
tttttgaccc tggagttccc cggtactggg gcatataaag tttgcaagac caaggggcct 17400
ctcttcccag tgatggccga ctaagccatc ttctgctaca tatgcagata gagacacgag 17460
ctctgggggt actagttagt tcatattgtt gttccaccta tagggtcgca gacccttca 17520
gctccttggg tactttgtct agctcctcca ctgggggctc tgtgttttat ctaatagatg 17580
actgtgagca tccacttctg tatttgacag gcactggcct agcgtcacat gagccagcta 17640
tatcagggtc ctttcagcaa aaccttgctg gcatgtgcaa tagtgtctgc gtttggtggt 17700
tgattatggg atggatccac tagttctaga gcggccgcca ccgcggtgga gctccagctt 17760
ttgttccctt tagtgagggt taattgcgcg cttggcgtaa tcatggtcat agctgtttcc 17820
tgtgtgaaat tgttatccgc tcacaattcc acacaacata cgagccggaa gcataaagtg 17880
taaagcctgg ggtgcctaat gagtgagcta actcacatta attgcgttgc gctcactgcc 17940
cgctttccag tcgggaaacc tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg 18000
gagaggcggt ttgcgtattg ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc 18060
ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaatac ggttatccac 18120
agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa 18180
ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgccccctg acgagcatca 18240
caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc 18300

gtttcccccc ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata 18360
cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta 18420
tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca 18480
gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga 18540
cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg 18600
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagga cagtatttgg 18660
tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg 18720
caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag 18780
aaaaaaagga tctcaagaag atcctttgat cttttctacg gggtctgacg ctcagtggaa 18840
cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat 18900
ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 18960
tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc 19020
atccatagtt gcctgactcc ccgtcgtgta gataactacg atacgggagg gcttaccatc 19080
tggccccagt gctgcaatga taccgcgaga cccacgctca ccggctccag atttatcagc 19140
aataaaccag ccagccggaa gggccgagcg cagaagtggt cctgcaactt tatccgcctc 19200
catccagtct attaattgtt gccgggaagc tagagtaagt agttcgccag ttaatagttt 19260
gcgcaacgtt gttgccattg ctacaggcat cgtggtgtca cgctcgtcgt ttggtatggc 19320
ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatccccca tgttgtgcaa 19380
aaaagcggtt agctccttcg gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt 19440
atcactcatg gttatggcag cactgcataa ttctcttact gtcatgccat ccgtaagatg 19500
cttttctgtg actggtgagt actcaaccaa gtcattctga gaatagtgta tgcggcgacc 19560
gagttgctct tgcccggcgt caatacggga taataccgcg ccacatagca gaactttaaa 19620
agtgctcatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct taccgctgtt 19680
gagatccagt tcgatgtaac ccactcgtgc acccaactga tcttcagcat cttttacttt 19740
caccagcgtt tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa agggaataag 19800
ggcgacacgg aaatgttgaa tactcatact cttccttttt caatattatt gaagcattta 19860
tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat 19920
aggggttccg cgcacatttc cccgaaaagt gccacctaaa ttgtaagcgt taatattttg 19980
ttaaaattcg cgttaaattt ttgttaaatc agctcatttt ttaaccaata ggccgaaatc 20040
ggcaaaatcc cttataaatc aaaagaatag accgagatag ggttgagtgt tgttccagtt 20100
tggaacaaga gtccactatt aaagaacgtg gactccaacg tcaaagggcg aaaaaccgtc 20160
tatcagggcg atggcccact acgtgaacca tcaccctaat caagtttttt ggggtcgagg 20220
tgccgtaaag cactaaatcg gaaccctaaa gggagccccc gatttagagc ttgacgggga 20280
aagccggcga acgtggcgag aaaggaaggg aagaaagcga aaggagcggg cgctagggcg 20340
ctggcaagtg tagcggtcac gctgcgcgta accaccacac ccgccgcgct taatgcgccg 20400
ctacagggcg cgtcccattc gccattcagg ctgcgcaact gttgggaagg gcgatcggtg 20460
cgggcctctt cgctattacg ccagctggcg aaaggggat gtgctgcaag gcgattaagt 20520
tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgagcgcgcg 20580
taatacgact cactataggg cgaattgggt accgggcccc ccc 20623

<210> 2
<211> 6708
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: R15/APPA +

intron plasmid with pBLCAT3 vector

<400> 2

ggatcccctt tgctatgtag tttttaatgg aaattacaac ccatagtgtg ttgataaata 60
gagagtcctg tttggtttaa gcaacctctg tttctcataa actccataaa aacaggaata 120
ctctttgttt ctagcataac caaaagattt agtgaattga aaacaatgtt cccttagagt 180
ataggtctaa taaccccgaa aatattacca tgatactgag catttgtaag tatctcatag 240
catgtagtat ccatagtcca tcaatgagag agacatttaa catgattttc attaatcagg 300
tggaaaagac atgacaacat tcacaggcac tgcacagaac atagtggtcc accttgcaca 360
tatttcacta aactaggttt atctattttg ttgctttctc taacatctct gcaatgaagc 420
aggtcaacag tgccacatat cctttactta acctaaggaa cacaaaaaat tttctacata 480
tatcctggtt agagagtgct taaaataagt tttccaagaa tggaaaagaa atgttctgac 540
ttaacaatta agacagtatt tatttaaagc aagaaatatg aggcacacaa gaaaatattt 600
tgggaagaaa ccatttggtg aacaatattt caaataaaaa tagacaaaca tagttaattg 660
taaaacatat gtttgaccag cccttctttt caataggctt aatgtgaata aaatgttaaa 720
gattctcttt gggtggctgc aaattgtcca cgaataagac aaaatataaa aataaggact 780
gagtctcaca aaatgaaaag gaaatatatt cagaaagaga atcttgagag aatgtgttgt 840
cacaaattaa agaaaacctg tggtgaatga catcctgagg cctgagctat tactgacatt 900
taagataaag gtaactgtat acatttgtcc cattgagggg acaagaaagc tgctctcatg 960
ttcagctcta taattcttgc cttaaacaac ttaaatagaa tgatttaaaa tatggagctg 1020
tccatggacc tttgaaatat aaaatagtca agcaacttat caaggaatta cagattcctt 1080
gatactaaca caggtaaatc ccacacgtgt tttgagacta catttgctgg gattttattg 1140
atgtaatagg tcacatgttt ttcgggccaa tgttgctgtt attcggttac ttcaagagaa 1200
tagtggcaac tgatgctatg tattctaggg gtttgaagtg atgtttcatg attgaaattt 1260
gtaaaagaat aacatcatca ttcttaacaa tagaacatat aaagtcacac agaagtgaca 1320
gtgtttaagc tgtactattg atcaaagaaa tttattacct tcagtttcaa tggaaataat 1380
tactgataat acaaacatgt gtgaacacac actaatccta tccaaatgca cagtgataca 1440
cagaaaatat tagcaagtag aatgcaatat ttatataacg attgtattta tcaatcaatt 1500
gtatgtatca atatatgggc tattttctta cacatgattt tattcaaatt tactctaatc 1560
attgttgaac catttagaaa aggcatactg gcaacttttc cttacctcat ccagctgggc 1620
aaaagtccca gtgtggagta aaggatgcaa gatttcctgc tctgttaagt ataaaataat 1680
agtatgaatt caaaggtgcc attcttctgc ttctagttat aaaggcagtg cttgcttctt 1740
ccagcacaga tctggatctc gaggagcttg gcgagatttt caggagctaa ggaagctaaa 1800
agccgccacc atgaaagcca tcttaatccc atttttatct cttctgattc cgttaacccc 1860
gcaatctgca ttcgctcaga gtgagccgga gctgaagctg gaaagtgtgg tgattgtcag 1920
tcgtcatggt gtgcgtgctc caaccaaggc cacgcaactg atgcaggatg tcaccccaga 1980
cgcatggcca acctggccgg taaaactggg ttggctgaca ccgcgcggtg gtgagctaat 2040
cgcctatctc ggacattacc aacgccagcg tctggtagcc gacggattgc tggcgaaaaa 2100
gggctgcccg cagtctggtc aggtcgcgat tattgctgat gtcgacgagc gtacccgtaa 2160
aacaggcgaa gccttcgccg ccgggctggc acctgactgt gcaataaccg tacataccca 2220
ggcagatacg tccagtcccg atccgttatt taatcctcta aaaactggcg tttgccaact 2280
ggataacgcg aacgtgactg acgcgatcct cagcagggca ggagggtcaa ttgctgactt 2340
taccgggcat cggcaaacgg cgtttcgcga actggaacgg gtgcttaatt ttccgcaatc 2400
aaacttgtgc cttaaacgtg agaaacagga cgaaagctgt tcattaacgc aggcattacc 2460
atcggaactc aaggtgagcg ccgacaatgt ctcattaacc ggtgcggtaa gcctcgcatc 2520
aatgctgacg gagatatttc tcctgcaaca agcacaggga atgccggagc cggggtgggg 2580
aaggatcacc gattcacacc agtggaacac cttgctaagt ttgcataacg cgcaatttta 2640
tttgctacaa cgcacgccag aggttgcccg cagccgcgcc accccgttat tagatttgat 2700

caagacagcg ttgacgcccc atccaccgca aaaacaggcg tatggtgtga cattacccac 2760
ttcagtgctg tttatcgccg gacacgatac taatctggca aatctcggcg gcgcactgga 2820
gctcaactgg acgcttcccg gtcagccgga taacacgccg ccaggtggtg aactggtgtt 2880
tgaacgctgg cgtcggctaa gcgataacag ccagtggatt caggtttcgc tggtcttcca 2940
gactttacag cagatgcgtg ataaaacgcc gctgtcatta aatacgccgc ccggagaggt 3000
gaaactgacc ctggcaggat gtgaagagcg aaatgcgcag ggcatgtgtt cgttggcagg 3060
ttttacgcaa atcgtgaatg aagcacgcat acccgcttgc agtttgtaag gtataaggca 3120
gttattggtg cccttaaacg cctggtgcta cgcctgaata agtgataata agcggatgaa 3180
tggcagaaat tcgccggatc tttgtgaagg aaccttactt ctgtggtgtg acataattgg 3240
acaaactacc tacagagatt taaagctcta aggtaaatat aaaattttta agtgtataat 3300
gtgttaaact actgattcta attgtttgtg tattttagat tccaacctat ggaactgatg 3360
aatgggagca gtggtggaat gcctttaatg aggaaaacct gttttgctca gaagaaatgc 3420
catctagtga tgatgaggct actgctgact ctcaacattc tactcctcca aaaaagaaga 3480
gaaaggtaga agaccccaag gactttcctt cagaattgct aagtttttg agtcatgctg 3540
tgtttagtaa tagaactctt gcttgctttg ctatttacac cacaaaggaa aaagctgcac 3600
tgctatacaa gaaaattatg gaaaaatatt ctgtaacctt tataagtagg cataacagtt 3660
ataatcataa catactgttt tttcttactc cacacaggca tagagtgtct gctattaata 3720
actatgctca aaaattgtgt acctttagct ttttaatttg taaaggggtt aataaggaat 3780
atttgatgta tagtgccttg actagagatc ataatcagcc ataccacatt tgtagaggtt 3840
ttacttgctt taaaaaacct cccacacctc cccctgaacc tgaaacataa aatgaatgca 3900
attgttgttg ttaacttgtt tattgcagct tataatggtt acaaataaag caatagcatc 3960
acaaatttca caaataaagc attttttttca ctgcattcta gttgtggttt gtccaaactc 4020
atcaatgtat cttatcatgt ctggatcgat ccccgggtac cgagctcgaa ttcgtaatca 4080
tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatacga 4140
gccggaagca taaagtgtaa agcctggggt gcctaatgag tgagctaact cacattaatt 4200
gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagct gcattaatga 4260
atcggccaac gcgcggggag aggcggtttg cgtattgggc gctcttccgc ttcctcgctc 4320
actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg 4380
gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc 4440
cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc 4500
cccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga 4560
ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc tgttccgacc 4620
ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcaa 4680
tgctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct gggctgtgtg 4740
cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc 4800
aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga 4860
gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta cggctacact 4920
agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt 4980
ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag 5040
cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg 5100
tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgag attatcaaaa 5160
aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata 5220
tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg 5280
atctgtctat ttcgttcatc catagttgcc tgactccccg tcgtgtagat aactacgata 5340
cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc acgctcaccg 5400
gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag aagtggtcct 5460
gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag agtaagtagt 5520
tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt ggtgtcacgc 5580

tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga 5640
tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt 5700
aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc tcttactgtc 5760
atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa 5820
tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca 5880
catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca 5940
aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc caactgatct 6000
tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc 6060
gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt cctttttcaa 6120
tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt 6180
tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc 6240
taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt 6300
cgtctcgcgc gtttcggtga tgacggtgaa aacctctgac acatgcagct cccggagacg 6360
gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg cgcgtcagcg 6420
ggtgttggcg ggtgtcgggg ctggcttaac tatgcggcat cagagcagat tgtactgaga 6480
gtgcaccata tgcggtgtga aataccgcac agatgcgtaa ggagaaaata ccgcatcagg 6540
cgccattcgc cattcaggct gcgcaactgt tgggaagggc gatcggtgcg ggcctcttcg 6600
ctattacgcc agctggcgaa agggggatgt gctgcaaggc gattaagttg ggtaacgcca 6660
gggttttccc agtcacgacg ttgtaaaacg acggccagtg ccaagctt 6708

<210> 3
<211> 4060
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: R15/APPA +
intron transgene

<400> 3
ggatcccctt tgctatgtag tttttaatgg aaattacaac ccatagtgtg ttgataaata 60
gagagtcctg tttggtttaa gcaacctctg tttctcataa actccataaa aacaggaata 120
ctctttgttt ctagcataac caaaagattt agtgaattga aaacaatgtt cccttagagt 180
ataggtctaa taaccccgaa aatattacca tgatactgag catttgtaag tatctcatag 240
catgtagtat ccatagtcca tcaatgagag agacatttaa catgattttc attaatcagg 300
tggaaaagac atgacaacat tcacaggcac tgcacagaac atagtggtcc accttgcaca 360
tatttcacta aactaggttt atctattttg ttgctttctc taacatctct gcaatgaagc 420
aggtcaacag tgccacatat cctttactta acctaaggaa cacaaaaaat tttctacata 480
tatcctggtt agagagtgct taaaataagt tttccaagaa tggaaaagaa atgttctgac 540
ttaacaatta agacagtatt tatttaaagc aagaaatatg aggcacacaa gaaaatattt 600
tgggaagaaa ccatttggtg aacaatattt caaataaaaa tagacaaaca tagttaattg 660
taaaacatat gtttgaccag ccctttcttt caataggctt aatgtgaata aaatgttaaa 720
gattctcttt gggtggctgc aaattgtcca cgaataagac aaaatataaa aataaggact 780
gagtctcaca aaatgaaaag gaaatatatt cagaaagaga atcttgagag aatgtgttgt 840
cacaaattaa agaaaacctg tggtgaatga catcctgagg cctgagctat tactgacatt 900
taagataaag gtaactgtat acatttgtcc cattgagggg acaagaaagc tgctctcatg 960
ttcagctcta taattcttgc cttaaacaac ttaaatagaa tgatttaaaa tatggagctg 1020

tccatggacc tttgaaatat aaaatagtca agcaacttat caaggaatta cagattcctt 1080
gatactaaca caggtaaatc ccacacgtgt tttgagacta catttgctgg gattttattg 1140
atgtaatagg tcacatgttt ttcgggccaa tgttgctgtt attcggttac ttcaagagaa 1200
tagtggcaac tgatgctatg tattctaggg gtttgaagtg atgtttcatg attgaaattt 1260
gtaaaagaat aacatcatca ttcttaacaa tagaacatat aaagtcacac agaagtgaca 1320
gtgtttaagc tgtactattg atcaaagaaa tttattacct tcagtttcaa tggaaataat 1380
tactgataat acaaacatgt gtgaacacac actaatccta tccaaatgca cagtgataca 1440
cagaaaatat tagcaagtag aatgcaatat ttatataacg attgtattta tcaatcaatt 1500
gtatgtatca atatatgggc tattttctta cacatgattt tattcaaatt tactctaatc 1560
attgttgaac catttagaaa aggcatactg gcaacttttc cttacctcat ccagctgggc 1620
aaaagtccca gtgtggagta aaggatgcaa gatttcctgc tctgttaagt ataaaataat 1680
agtatgaatt caaaggtgcc attcttctgc ttctagttat aaaggcagtg cttgcttctt 1740
ccagcacaga tctggatctc gaggagcttg gcgagatttt caggagctaa ggaagctaaa 1800
agccgccacc atgaaagcca tcttaatccc atttttatct cttctgattc cgttaacccc 1860
gcaatctgca ttcgctcaga gtgagccgga gctgaagctg gaaagtgtgg tgattgtcag 1920
tcgtcatggt gtgcgtgctc caaccaaggc cacgcaactg atgcaggatg tcaccccaga 1980
cgcatggcca acctggccgg taaaactggg ttggctgaca ccgcgcggtg gtgagctaat 2040
cgcctatctc ggacattacc aacgccagcg tctggtagcc gacggattgc tggcgaaaaa 2100
gggctgcccg cagtctggtc aggtcgcgat tattgctgat gtcgacgagc gtacccgtaa 2160
aacaggcgaa gccttcgccg ccgggctggc acctgactgt gcaataaccg tacataccca 2220
ggcagatacg tccagtcccg atccgttatt taatcctcta aaaactggcg tttgccaact 2280
ggataacgcg aacgtgactg acgcgatcct cagcagggca ggagggtcaa ttgctgactt 2340
taccgggcat cggcaaacgg cgtttcgcga actggaacgg gtgcttaatt ttccgcaatc 2400
aaacttgtgc cttaaacgtg agaaacagga cgaaagctgt tcattaacgc aggcattacc 2460
atcggaactc aaggtgagcg ccgacaatgt ctcattaacc ggtgcggtaa gcctcgcatc 2520
aatgctgacg gagatatttc tcctgcaaca agcacaggga atgccggagc cggggtgggg 2580
aaggatcacc gattcacacc agtggaacac cttgctaagt ttgcataacg cgcaatttta 2640
tttgctacaa cgcacgccag aggttgcccg cagccgcgcc accccgttat tagatttgat 2700
caagacagcg ttgacgcccc atccaccgca aaaacaggcg tatggtgtga cattacccac 2760
ttcagtgctg tttatcgccg gacacgatac taatctggca aatctcggcg gcgcactgga 2820
gctcaactgg acgcttcccg gtcagccgga taacacgccg ccaggtggtg aactggtgtt 2880
tgaacgctgg cgtcggctaa gcgataacag ccagtggatt caggtttcgc tggtcttcca 2940
gactttacag cagatgcgtg ataaaacgcc gctgtcatta aatacgccgc ccggagaggt 3000
gaaactgacc ctggcaggat gtgaagagcg aaatgcgcag ggcatgtgtt cgttggcagg 3060
ttttacgcaa atcgtgaatg aagcacgcat acccgcttgc agtttgtaag gtataaggca 3120
gttattggtg cccttaaacg cctggtgcta cgcctgaata agtgataata agcggatgaa 3180
tggcagaaat tcgccggatc tttgtgaagg aaccttactt ctgtggtgtg acataattgg 3240
acaaactacc tacagagatt taaagctcta aggtaaatat aaaattttta agtgtataat 3300
gtgttaaact actgattcta attgtttgtg tattttagat tccaacctat ggaactgatg 3360
aatgggagca gtggtggaat gcctttaatg aggaaaacct gttttgctca gaagaaatgc 3420
catctagtga tgatgaggct actgctgact ctcaacattc tactcctcca aaaaagaaga 3480
gaaaggtaga agaccccaag gactttcctt cagaattgct aagtttttg agtcatgctg 3540
tgtttagtaa tagaactctt gcttgctttg ctatttacac cacaaaggaa aaagctgcac 3600
tgctatacaa gaaaattatg gaaaaatatt ctgtaacctt tataagtagg cataacagtt 3660
ataatcataa catactgttt tttcttactc cacacaggca tagagtgtct gctattaata 3720
actatgctca aaaattgtgt acctttagct ttttaatttg taaaggggtt aataaggaat 3780
atttgatgta tagtgccttg actagagatc ataatcagcc ataccacatt tgtagaggtt 3840
ttacttgctt taaaaaacct cccacacctc cccctgaacc tgaaacataa aatgaatgca 3900

attgttgttg ttaacttgtt tattgcagct tataatggtt acaaataaag caatagcatc 3960
acaaatttca caaataaagc attttttca ctgcattcta gttgtggttt gtccaaactc 4020
atcaatgtat cttatcatgt ctggatcgat ccccgggtac 4060

<210> 4
<211> 6116
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: R15/APPA
plasmid with pBLCAT3 vector

<400> 4
ggatcccctt tgctatgtag tttttaatgg aaattacaac ccatagtgtg ttgataaata 60
gagagtcctg tttggtttaa gcaacctctg tttctcataa actccataaa aacaggaata 120
ctctttgttt ctagcataac caaaagattt agtgaattga aaacaatgtt cccttagagt 180
ataggtctaa taaccccgaa aatattacca tgatactgag catttgtaag tatctcatag 240
catgtagtat ccatagtcca tcaatgagag agacatttaa catgattttc attaatcagg 300
tggaaaagac atgacaacat tcacaggcac tgcacagaac atagtggtcc accttgcaca 360
tatttcacta aactaggttt atctattttg ttgctttctc taacatctct gcaatgaagc 420
aggtcaacag tgccacatat cctttactta acctaaggaa cacaaaaaat tttctacata 480
tatcctggtt agagagtgct taaaataagt tttccaagaa tggaaaagaa atgttctgac 540
ttaacaatta agacagtatt tatttaaagc aagaaatatg aggcacacaa gaaaatattt 600
tgggaagaaa ccatttggtg aacaatattt caaataaaaa tagacaaaca tagttaattg 660
taaaacatat gtttgaccag ccctttcttt caataggctt aatgtgaata aaatgttaaa 720
gattctcttt gggtggctgc aaattgtcca cgaataagac aaaatataaa aataaggact 780
gagtctcaca aaatgaaaag gaaatatatt cagaaagaga atcttgagag aatgtgttgt 840
cacaaattaa agaaaacctg tggtgaatga catcctgagg cctgagctat tactgacatt 900
taagataaag gtaactgtat acatttgtcc cattgagggg acaagaaagc tgctctcatg 960
ttcagctcta taattcttgc cttaaacaac ttaaatagaa tgatttaaaa tatggagctg 1020
tccatggacc tttgaaatat aaaatagtca agcaacttat caaggaatta cagattcctt 1080
gatactaaca caggtaaatc ccacacgtgt tttgagacta catttgctgg gattttattg 1140
atgtaatagg tcacatgttt ttcgggccaa tgttgctgtt attcggttac ttcaagagaa 1200
tagtggcaac tgatgctatg tattctaggg gtttgaagtg atgtttcatg attgaaattt 1260
gtaaaagaat aacatcatca ttcttaacaa tagaacatat aaagtcacac agaagtgaca 1320
gtgtttaagc tgtactattg atcaaagaaa tttattacct tcagtttcaa tggaaataat 1380
tactgataat acaaacatgt gtgaacacac actaatccta tccaaatgca cagtgataca 1440
cagaaaatat tagcaagtag aatgcaatat ttatataacg attgtattta tcaatcaatt 1500
gtatgtatca atatatgggc tattttctta cacatgattt tattcaaatt tactctaatc 1560
attgttgaac catttagaaa aggcatactg gcaacttttc cttacctcat ccagctgggc 1620
aaaagtccca gtgtggagta aaggatgcaa gatttcctgc tctgttaagt ataaaataat 1680
agtatgaatt caaaggtgcc attcttctgc ttctagttat aaaggcagtg cttgcttctt 1740
ccagcacaga tctggatctc gaggagcttg gcgagatttt caggagctaa ggaagctaaa 1800
agccgccacc atgaaagcca tcttaatccc atttttatct cttctgattc cgttaacccc 1860
gcaatctgca ttcgctcaga gtgagccgga gctgaagctg gaaagtgtgg tgattgtcag 1920
tcgtcatggt gtgcgtgctc caaccaaggc cacgcaactg atgcaggatg tcaccccaga 1980

cgcatggcca acctggccgg taaaactggg ttggctgaca ccgcgcggtg gtgagctaat 2040
cgcctatctc ggacattacc aacgccagcg tctggtagcc gacggattgc tggcgaaaaa 2100
gggctgcccg cagtctggtc aggtcgcgat tattgctgat gtcgacgagc gtacccgtaa 2160
aacaggcgaa gccttcgccg ccgggctggc acctgactgt gcaataaccg tacataccca 2220
ggcagatacg tccagtcccg atccgttatt taatcctcta aaaactggcg tttgccaact 2280
ggataacgcg aacgtgactg acgcgatcct cagcagggca ggagggtcaa ttgctgactt 2340
taccgggcat cggcaaacgg cgtttcgcga actggaacgg gtgcttaatt ttccgcaatc 2400
aaacttgtgc cttaaacgtg agaaacagga cgaaagctgt tcattaacgc aggcattacc 2460
atcggaactc aaggtgagcg ccgacaatgt ctcattaacc ggtgcggtaa gcctcgcatc 2520
aatgctgacg gagatatttc tcctgcaaca agcacaggga atgccggagc cggggtgggg 2580
aaggatcacc gattcacacc agtggaacac cttgctaagt ttgcataacg cgcaatttta 2640
tttgctacaa cgcacgccag aggttgcccg cagccgcgcc accccgttat tagatttgat 2700
caagacagcg ttgacgcccc atccaccgca aaaacaggcg tatggtgtga cattacccac 2760
ttcagtgctg tttatcgccg gacacgatac taatctggca aatctcggcg gcgcactgga 2820
gctcaactgg acgcttcccg gtcagccgga taacacgccg ccaggtggtg aactggtgtt 2880
tgaacgctgg cgtcggctaa gcgataacag ccagtggatt caggtttcgc tggtcttcca 2940
gactttacag cagatgcgtg ataaaacgcc gctgtcatta aatacgccgc ccggagaggt 3000
gaaactgacc ctggcaggat gtgaagagcg aaatgcgcag ggcatgtgtt cgttggcagg 3060
ttttacgcaa atcgtgaatg aagcacgcat acccgcttgc agtttgtaag gtataaggca 3120
gttattggtg cccttaaacg cctggtgcta cgcctgaata agtgataata agcggatgaa 3180
tggcagaaat tcgccggatc tttgtgaagg aaccttactt ctgtggtgtg acataattgg 3240
acaaactacc tacagagatt taaaaaacct cccacacctc cccctgaacc tgaaacataa 3300
aatgaatgca attgttgttg ttaacttgtt tattgcagct tataatggtt acaaataaag 3360
caatagcatc acaaatttca caaataaagc attttttttca ctgcattcta gttgtggttt 3420
gtccaaactc atcaatgtat cttatcatgt ctggatcgat ccccgggtac cgagctcgaa 3480
ttcgtaatca tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca 3540
caacatacga gccggaagca taaagtgtaa agcctggggt gcctaatgag tgagctaact 3600
cacattaatt gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagct 3660
gcattaatga atcggccaac gcgcggggag aggcggtttg cgtattgggc gctcttccgc 3720
ttcctcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca 3780
ctcaaaggcg gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg 3840
agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca 3900
taggctccgc ccccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa 3960
cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc 4020
tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc 4080
gctttctcaa tgctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct 4140
gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg 4200
tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag 4260
gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta 4320
cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg 4380
aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt 4440
tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt 4500
ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgag 4560
attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat 4620
ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc 4680
tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactccccg tcgtgtagat 4740
aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc 4800
acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag 4860

aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag 4920
agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt 4980
ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg 5040
agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt 5100
tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc 5160
tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc 5220
attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa 5280
taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg 5340
aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc 5400
caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag 5460
gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt 5520
cctttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt 5580
tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc 5640
acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 5700
gaggcccttt cgtctcgcgc gtttcggtga tgacggtgaa aacctctgac acatgcagct 5760
cccggagacg gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg 5820
cgcgtcagcg ggtgttggcg ggtgtcgggg ctggcttaac tatgcggcat cagagcagat 5880
tgtactgaga gtgcaccata tgcggtgtga aataccgcac agatgcgtaa ggagaaaata 5940
ccgcatcagg cgccattcgc cattcaggct gcgcaactgt tgggaagggc gatcggtgcg 6000
ggcctcttcg ctattacgcc agctggcgaa aggggatgt gctgcaaggc gattaagttg 6060
ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg ccaagc 6116

<210> 5
<211> 3470
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: R15/APPA
transgene

<400> 5
ggatcccctt tgctatgtag tttttaatgg aaattacaac ccatagtgtg ttgataaata 60
gagagtcctg tttggtttaa gcaacctctg tttctcataa actccataaa aacaggaata 120
ctctttgttt ctagcataac caaaagattt agtgaattga aaacaatgtt cccttagagt 180
ataggtctaa taaccccgaa aatattacca tgatactgag catttgtaag tatctcatag 240
catgtagtat ccatagtcca tcaatgagag agacatttaa catgattttc attaatcagg 300
tggaaaagac atgacaacat tcacaggcac tgcacagaac atagtggtcc accttgcaca 360
tatttcacta aactaggttt atctattttg ttgctttctc taacatctct gcaatgaagc 420
aggtcaacag tgccacatat cctttactta acctaaggaa cacaaaaaat tttctacata 480
tatcctggtt agagagtgct taaaataagt tttccaagaa tggaaaagaa atgttctgac 540
ttaacaatta agacagtatt tatttaaagc aagaaatatg aggcacacaa gaaaatattt 600
tgggaagaaa ccatttggtg aacaatattt caaataaaaa tagacaaaca tagttaattg 660
taaaacatat gtttgaccag cccttctttt caataggctt aatgtgaata aaatgttaaa 720
gattctcttt gggtggctgc aaattgtcca cgaataagac aaaatataaa aataaggact 780
gagtctcaca aaatgaaaag gaaatatatt cagaaagaga atcttgagag aatgtgttgt 840
cacaaattaa agaaaacctg tggtgaatga catcctgagg cctgagctat tactgacatt 900

taagataaag gtaactgtat acatttgtcc cattgagggg acaagaaagc tgctctcatg 960
ttcagctcta taattcttgc cttaaacaac ttaaatagaa tgatttaaaa tatggagctg 1020
tccatggacc tttgaaatat aaaatagtca agcaacttat caaggaatta cagattcctt 1080
gatactaaca caggtaaatc ccacacgtgt tttgagacta catttgctgg gattttattg 1140
atgtaatagg tcacatgttt ttcgggccaa tgttgctgtt attcggttac ttcaagagaa 1200
tagtggcaac tgatgctatg tattctaggg gtttgaagtg atgtttcatg attgaaattt 1260
gtaaaagaat aacatcatca ttcttaacaa tagaacatat aaagtcacac agaagtgaca 1320
gtgtttaagc tgtactattg atcaaagaaa tttattacct tcagtttcaa tggaaataat 1380
tactgataat acaaacatgt gtgaacacac actaatccta tccaaatgca cagtgataca 1440
cagaaaatat tagcaagtag aatgcaatat ttatataacg attgtattta tcaatcaatt 1500
gtatgtatca atatatgggc tattttctta cacatgattt tattcaaatt tactctaatc 1560
attgttgaac catttagaaa aggcatactg gcaacttttc cttacctcat ccagctgggc 1620
aaaagtccca gtgtggagta aaggatgcaa gatttcctgc tctgttaagt ataaaataat 1680
agtatgaatt caaaggtgcc attcttctgc ttctagttat aaaggcagtg cttgcttctt 1740
ccagcacaga tctggatctc gaggagcttg gcgagatttt caggagctaa ggaagctaaa 1800
agccgccacc atgaaagcca tcttaatccc atttttatct cttctgattc cgttaacccc 1860
gcaatctgca ttcgctcaga gtgagccgga gctgaagctg gaaagtgtgg tgattgtcag 1920
tcgtcatggt gtgcgtgctc caaccaaggc cacgcaactg atgcaggatg tcaccccaga 1980
cgcatggcca acctggccgg taaaactggg ttggctgaca ccgcgcggtg gtgagctaat 2040
cgcctatctc ggacattacc aacgccagcg tctggtagcc gacggattgc tggcgaaaaa 2100
gggctgcccg cagtctggtc aggtcgcgat tattgctgat gtcgacgagc gtacccgtaa 2160
aacaggcgaa gccttcgccg ccgggctggc acctgactgt gcaataaccg tacataccca 2220
ggcagatacg tccagtcccg atccgttatt taatcctcta aaaactggcg tttgccaact 2280
ggataacgcg aacgtgactg acgcgatcct cagcagggca ggagggtcaa ttgctgactt 2340
taccgggcat cggcaaacgg cgtttcgcga actggaacgg gtgcttaatt ttccgcaatc 2400
aaacttgtgc cttaaacgtg agaaacagga cgaaagctgt tcattaacgc aggcattacc 2460
atcggaactc aaggtgagcg ccgacaatgt ctcattaacc ggtgcggtaa gcctcgcatc 2520
aatgctgacg gagatatttc tcctgcaaca agcacaggga atgccggagc cggggtgggg 2580
aaggatcacc gattcacacc agtggaacac cttgctaagt ttgcataacg cgcaatttta 2640
tttgctacaa cgcacgccag aggttgcccg cagccgcgcc accccgttat tagatttgat 2700
caagacagcg ttgacgcccc atccaccgca aaaacaggcg tatggtgtga cattacccac 2760
ttcagtgctg tttatcgccg gacacgatac taatctggca aatctcggcg gcgcactgga 2820
gctcaactgg acgcttcccg gtcagccgga taacacgccg ccaggtggtg aactggtgtt 2880
tgaacgctgg cgtcggctaa gcgataacag ccagtggatt caggtttcgc tggtcttcca 2940
gactttacag cagatgcgtg ataaaacgcc gctgtcatta aatacgccgc ccggagaggt 3000
gaaactgacc ctggcaggat gtgaagagcg aaatgcgcag ggcatgtgtt cgttggcagg 3060
ttttacgcaa atcgtgaatg aagcacgcat acccgcttgc agtttgtaag gtataaggca 3120
gttattggtg cccttaaacg cctggtgcta cgcctgaata agtgataata agcggatgaa 3180
tggcagaaat tcgccggatc tttgtgaagg aaccttactt ctgtggtgtg acataattgg 3240
acaaactacc tacagagatt taaaaaacct cccacacctc cccctgaacc tgaaacataa 3300
aatgaatgca attgttgttg ttaacttgtt tattgcagct tataatggtt acaaataaag 3360
caatagcatc acaaatttca caaataaagc attttttttca ctgcattcta gttgtggttt 3420
gtccaaactc atcaatgtat cttatcatgt ctggatcgat ccccgggtac 3470

<210> 6
<211> 5421
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: SV40/APPA + intron plasmid

<400> 6
cgagattttc aggagctaag gaagctaaaa gccgccacca tgaaagccat cttaatccca 60
tttttatctc ttctgattcc gttaaccccg caatctgcat tcgctcagag tgagccggag 120
ctgaagctgg aaagtgtggt gattgtcagt cgtcatggtg tgcgtgctcc aaccaaggcc 180
acgcaactga tgcaggatgt caccccagac gcatggccaa cctggccggt aaaactgggt 240
tggctgacac cgcgnggtgg tgagctaatc gcctatctcg gacattacca acgccagcgt 300
ctggtagccg acggattgct ggcgaaaaag ggctgcccgc agtctggtca ggtcgcgatt 360
attgctgatg tcgacgagcg tacccgtaaa acaggcgaag ccttcgccgc cgggctggca 420
cctgactgtg caataaccgt acatacccag gcagatacgt ccagtcccga tccgttattt 480
aatcctctaa aaactggcgt ttgccaactg gataacgcga acgtgactga cgcgatcctc 540
agcagggcag gagggtcaat tgctgacttt accgggcatc ggcaaacggc gtttcgcgaa 600
ctggaacggg tgcttaattt tccgcaatca aacttgtgcc ttaaacgtga gaaacaggac 660
gaaagctgtt cattaacgca ggcattacca tcggaactca aggtgagcgc cgacaatgtc 720
tcattaaccg gtgcggtaag cctcgcatca atgctgacgg agatatttct cctgcaacaa 780
gcacagggaa tgccggagcc ggggtgggga aggatcaccg attcacacca gtggaacacc 840
ttgctaagtt tgcataacgc gcaatttttat ttgctacaac gcacgccaga ggttgcccgc 900
agccgcgcca ccccgttatt agatttgatc aagacagcgt tgacgcccca ccaccgcaaa 960
aacaggcgta tggtgtgaca ttacccactt cagtgctgtt tatcgccgga cacgatacta 1020
atctggcaaa tctcggcggc gcactggagc tcaactggac gcttcccggt cagccggata 1080
acacgccgcc aggtggtgaa ctggtgtttg aacgctggcg tcggctaagc gataacagcc 1140
agtggattca ggtttcgctg gtcttccaga ctttacagca gatgcgtgat aaaacgccgc 1200
tgtcattaaa tacgccgccc ggagaggtga aactgaccct ggcaggatgt gaagagcgaa 1260
atgcgcaggg catgtgttcg ttggcaggtt ttacgcaaat cgtgaatgaa gcacgcatac 1320
ccgcttgcag tttgtaaggc agttattggt gcccttaaac gcctggtgct acgcctgaat 1380
aagtgataat aagcggatga atggcagaaa ttcgccggat ctttgtgaag gaaccttact 1440
tctgtggtgt gacataattg gacaaactac ctacagagat ttaaagctct aaggtaaata 1500
taaaattttt aagtgtataa tgtgttaaac tactgattct aattgtttgt gtattttaga 1560
ttccaaccta tggaactgat gaatgggagc agtggtggaa tgcctttaat gaggaaaacc 1620
tgttttgctc agaagaaatg ccatctagtg atgatgaggc tactgctgac tctcaacatt 1680
ctactcctcc aaaaaagaag agaaaggtag aagaccccaa ggactttcct tcagaattgc 1740
taagtttttt gagtcatgct gtgtttagta atagaactct tgcttgcttt gctatttaca 1800
ccacaaagga aaaagctgca ctgctataca agaaaattat ggaaaaatat tctgtaacct 1860
ttataagtag gcataacagt tataatcata acatactgtt ttttcttact ccacacaggc 1920
atagagtgtc tgctattaat aactatgctc aaaaattgtg tacctttagc tttttaattt 1980
gtaaaggggt taataaggaa tatttgatgt atagtgcctt gactagagat cataatcagc 2040
cataccacat ttgtagaggt tttacttgct ttaaaaaacc tcccacacct ccccctgaac 2100
ctgaaacata aaatgaatgc aattgttgtt gttaacttgt ttattgcagc ttataatggt 2160
tacaaataaa gcaatagcat cacaaatttc acaaataaag catttttttc actgcattct 2220
agttgtggtt tgtccaaact catcaatgta tcttatcatg tctggatcga tccccgggta 2280
ccgagctcga attcgtaatc atggtcatag ctgtttcctg tgtgaaattg ttatccgctc 2340
acaattccac acaacatacg agccggaagc ataaagtgta aagcctgggg tgcctaatga 2400
gtgagctaac tcacattaat tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg 2460

tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg 2520
cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg 2580
gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga 2640
aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg 2700
gcgtttttcc ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag 2760
aggtggcgaa acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc 2820
gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg 2880
ggaagcgtgg cgctttctca atgctcacgc tgtaggtatc tcagttcggt gtaggtcgtt 2940
cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc 3000
ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc 3060
actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg 3120
tggcctaact acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca 3180
gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc 3240
ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat 3300
cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt 3360
ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt 3420
tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc 3480
agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc 3540
gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata 3600
ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg 3660
gccgagcgca gaagtggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc 3720
cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct 3780
acaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc cggttcccaa 3840
cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctccttcggt 3900
cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca 3960
ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac 4020
tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca 4080
atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt 4140
tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc 4200
actcgtgcac ccaactgatc ttcagcatct tttactttca ccagcgtttc tgggtgagca 4260
aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata 4320
ctcatactct tccttttca atattattga agcatttatc agggttattg tctcatgagc 4380
ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc 4440
cgaaaagtgc cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat 4500
aggcgtatca cgaggccctt tcgtctcgcg cgtttcggtg atgacggtga aaacctctga 4560
cacatgcagc tcccggagac ggtcacagct tgtctgtaag cggatgccgg gagcagacaa 4620
gcccgtcagg gcgcgtcagc gggtgttggc gggtgtcggg gctggcttaa ctatgcggca 4680
tcagagcaga ttgtactgag agtgcaccat atgcggtgtg aaataccgca cagatgcgta 4740
aggagaaaat accgcatcag gcgccattcg ccattcaggc tgcgcaactg ttgggaaggg 4800
cgatcggtgc gggcctcttc gctattacgc cagctggcga aaggggatg tgctgcaagg 4860
cgattaagtt gggtaacgcc agggttttcc cagtcacgac gttgtaaaac gacggccagt 4920
gccaagcttt acactttatg cttccggctc gtatgttgtg tggaattgtg agcggataac 4980
aatttcacac aggaaacagc tatgaccatg attacgaatt cggcgcagca ccatggcctg 5040
aaataacctc tgaaagagga acttggttag gtaccttctg aggcggaaag aaccagctgt 5100
ggaatgtgtg tcagttaggg tgtggaaagt ccccaggctc cccagcaggc agaagtatgc 5160
aaagcatgca tctcaattag tcagcaacca ggtgtggaaa gtccccaggc tccccagcag 5220
gcagaagtat gcaaagcatg catctcaatt agtcagcaac catagtcccg cccctaactc 5280
cgcccatccc gcccctaact ccgcccagtt ccgcccattc tccgccccat ggctgactaa 5340

tttttttat ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt 5400
gaggaggctc gaggagcttg g 5421

<210> 7
<211> 17732
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Lama2/APPA transgene

<400> 7
tcgagagtat ctttgtcagc tgtgcctcca acaaaggggt actgttgccc acatagaaag 60
atctaaacta attaattaat ccctcacccg caaatctttc agtcactaag ttagcacgat 120
tgttgaacaa gttctccaaa ggagagatac agatgagtgc gtatagggtg gacctggctg 180
ctgaggagac acctgcatct gactaagaag agccacggtg ttagttgaat ggtgtggagt 240
agggtggttc tgtgggacag tagaaaatcg agaggcatgt gccgtttagt gaactgatgg 300
aagctacccc aaacgacaga gattgtcagt caggccaatc cgtttcgagt ttgatgggca 360
gccggacagt gagacagaca cacctactca gttggaggaa ggatgagaac aatggccagc 420
agggattgag agaccctgac aggcgcaagg ccctaacaca cacacctacc acctcacttg 480
acaaagctgc caaagaccaa agacttgttc tccattagaa atgacagctg gcttgacccg 540
acagcataat aagcagagtg tactctgatt ggagaacttt aatgtgtttc attcagtatt 600
ataaaaggac agtattacag attttgttgt acactgctgt tacatgtggg gcagtgtgtc 660
tttaagtagg gtaaagtact ctttaaaaat gggtcctaga tatttttcc tttaactcaa 720
gtctcttact gtttaaatga tttttatttt gtttaatatg gaggaaaaag aagcgtaaat 780
ggacaatata tatttagaga aagatggtta gctgtcagaa aaatatgcaa atcaaaatca 840
caccaagact gcagcacacc cctgtcagat ggctgtgatc aagaaaataa atgacaatga 900
gtggtggtga agatgtacta aagggaaaca cacacacaca cacacacaca cacacacaca 960
cacactggag caaccactgt ggaaatcagt atgaatggtc ctcaaaaacc tgaagataga 1020
gcggggcgtg gtggcataca cttttattcc cagcactggg gaggcagagg caggtggatc 1080
tctgagttcc aggccagcct ggtctatagc acaggttcta ggacagccag ggctacacag 1140
aaaaaccctg ccttgattaa accaaaccaa accaaaccaa accaaaccaa accaaaccaa 1200
accaaaccaa accaaaccag accaaaccaa aacactgaag atagaacttc agtattccat 1260
tcctagatat atacccaatg gagactaagt cagcaagaca cctgcacagc catgttcact 1320
actacactgt tcaccacagc caggctgtgg aaccagcctg agtgtccatg ataaatgaat 1380
ggataggtaa ctttcaaggt aaatggactc tgctgtgtac atgcctcaca ttctgtttat 1440
tcatttttct ttatgaggtg tccattcagg agtcacatgg tagttctatt ttcagtcttc 1500
tgaagatact acactggtcc ccacagttta cacttttatc agcagtgaat aagggttcct 1560
ctatccttac catcatttgt tgtaattttt cttgatgacc ctctttctga cagggatagg 1620
atgtaatatc agtgtgagga agtacaactt gttttctaag tatttattgg ccccttgcat 1680
ttcttctttt gaaaactgtc ggttcctgac atctgctcag gtattcattg gatgttgttt 1740
ctttggtgtt tgagttctta tgaattctag atgttaaatc cctgcctgtg gttctctccc 1800
attctgtagg ctgcctcctc accctggcaa ttgttgtcct tgttttgcag aaacttttga 1860
cttcatggaa tctcatttgt cagttttccc tcctctgcta tagcctgagc taatgcactg 1920
gtttttacag agccctggtc tatgccttta tcctcctctg gcagcttcgg agtttcattt 1980
cttacattta gatctttgat ccactttgaa caagttttgg agcagggtga gagatacgaa 2040

tctagttcca ttcttccata tgtgatccta gtttacatag catcgttggt tgaagaggtt 2100
ttattttatt tttaaataat gtgtcataaa aaacgaggtg gttgtagcag tgtggatttg 2160
tttctttgtc ctttgatcta caggtcttgt tttgtgtcag tctcatgatg ttttattgct 2220
atggctctgt catacagtct gaggtcaggt attgtgatat accttcagta ttgctccctc 2280
agactcaggt ttgctttggc caggagtcat cttactcagt gctcttagag ctcccccagc 2340
atgtagctgc tactattctt agttgataaa tcaggaaact ggggctcaga gagattaact 2400
gtcttgaact acttctgggg aggtgaaacg tggagacact aaactgtgtt taccctgtac 2460
tgctccagta gctgtcgggt gctgggctac agcaaagcac ctatactata tattactcag 2520
gaggtggaaa aactcagcct cccttggggt tcccaagctc ccaggtgtcc agtcactgct 2580
ggaaacctca tggagtctga aaggaagggt tgagggtaca tggggcagcg atgaggagcc 2640
tggggctggg atctcccaaa cacctggata tccagatgcc actgggtcag ggggagttgg 2700
gaacagagtt gggatgtcca tggacctgtg acaaggccag ggccaggggg aggataactc 2760
tggctttact aatttgcgaa agtccttagc ttagcagcag ttgtctggga gcacagaggg 2820
gccttctgta agaggctcag gcagtgccgc tctgtaggcg aaggtcttct ccatgttccc 2880
catggtggtt cttgatgaaa gagacagtcc ttggctccaa actggtttat tgattgttca 2940
ttgtggaaaa tgggtgcaca ccaccttctc agggtggacc agagatcaaa taccttttgc 3000
agggaggaat atctgggaag ggacgcttac tggctaaacc ctcagggcct ctagatacat 3060
cattagcatg gagaactctg ttctgggcta catgaccaca ggccacattt ccacaagcca 3120
catgtgggaa gtgtggcaca tgttctaggc caggaatctg gtagggagcg tggagccacc 3180
taccatccca ggtgggtgcc tgggtgccag ggaccctgaa cccgctcaac cttaccaagt 3240
ttcctggcag ggtccactgt cctacacaga agctggagga ggtgtgaggg ttgtgtcttt 3300
gtggaatgtc ccatgctgct tggggctcag tttctccacc tgtacctcat tggtttgggt 3360
ataaaaagtg gggatacttt attattctct gactcggtcc tgaggaaaaa gcatcgtggc 3420
agtccaggaa ccacaccctg aggttcctgc actgaaggga ctccctaagt ctctggagtc 3480
tctcccctte acagagctgc caaagtctag gttcttttga ggataacaga gccatgcttg 3540
gtaagcagac aacagcattt gtttactcaa ccttcttttg tcagctccct cttcataaac 3600
aagttgagac accatgctgg cttgaggaag acttctaaag ccagacaact gtgcaaggaa 3660
gaagaagaag gggcaagtgg agttagcctg gatgtagccc tcaaagtctc cagagaccag 3720
ccatgaaggc tcaagtggag ggcaagacct gcagcagcca agcatctggc aggagaggat 3780
cctgggaacc cctctaccat gacacacatt cttcctgcag gtcacactta ataggccatt 3840
tcttatttgg atctatcatg gtgttctgtg cgagattaat gaggtgttat gctgcgaaca 3900
gaaagttata taaaaacaag tccccccccc ttgtcactgc tgctaagaat gtagcagaaa 3960
ttgtctcaag tgtctctcta atcagaaaca ataaaggtct ccttggattc aagccctcca 4020
gtttcctcct tccttgctga gccttggaca cccatacaaa cctcctggat gctacagctc 4080
tgggcagaga ctccaaggtg gggagagact gatggtacaa aagcaaaata cttgtttggg 4140
ggtacaccca ctcctctgcc tgtgtggttc ctgcagtcag tcctgcagac aggccctcag 4200
tgggtcttcc atgggcaaca cgcagaggga ggcaatggat gggaataccc acaccctggt 4260
tagtttaccc cggccatgct ctctgctctt catccctcct ctgccctctg ccacggcttt 4320
ctctgcagga atcatatctt catattggcc cacaggtgtt ctcctcaccc tagctatgat 4380
gtttacttta gagtgacctt agcagggctg gtgggaatga gttctagaag gctcacggag 4440
atgctaggga agaaacgtct tctaactact gaggttacta agttcctggt ggttgtctct 4500
gcctttccct tgttaaagtc accttgaagt tagtgcagaa gaaatcagag cccagtcaca 4560
gagtaaatat ggtcctgaag atttcctttg agtgcccaga atccatgaca tttcaagagc 4620
cctctttgta ccttaagtca tttggggttg tatcttctgc ttgatgtatg tgtgtgtgtt 4680
tatcaaagag tgagatggtt acataagagg tgctctaaag gacagagagg atttgcaatt 4740
gtggcatgtg acatcctcag gccttgctct ggtgccagga ggaactgatg cagaaaagag 4800
taagaggtca tttcctggag gctgtcacta tagaggagat cttacagtgc attccctcct 4860
ccaggccctg cctgaggata gacatgtgct gactgcaact gaaacagagg cttgggatgg 4920

agagttaggt tcacagaagg gagggtggga gatggatgct tgctgggttc tgggtctcat 4980
caccagctcc tgaccacccg gtcagcccat gtgcttattc catagctttc ttttgctatg 5040
tttactcagt gtggtgtttg ttgggaccca gcagaagcca gtcccaggct gacagctgtg 5100
gatacacagg gcagcatgag ggtcctcagc ctgaagcagt caggctggca gaagagaaag 5160
accagcacac attccttcaa ccaactatgt cttgaaaaac aaacatatta tatcacatat 5220
attgcattta tgagacagct aaaatgtact cgggtagcat gactccaggt ggggatatct 5280
gcaagtgcca tgagtggcag agggacagcc aatgtgaggc aagaaggaat tctggctcaa 5340
cacagcttag ctccctggtg ttggttcaaa ctttgagagt ttgaccacaa gcactttatt 5400
tttgacatat ttaaacagag cacaactttg ggaaaaagtt ttcttatgaa aattatcaca 5460
ataaagctta aggcatgact acattaaaat gcctttgcaa agtatatgtg ccctcttcca 5520
caagaatggt tctattgact gagaaataat gttcaggata aagatccagg aagaaaagat 5580
cagggataag taaaatacta aactcttttg caaagtacat agaccctctt tcataacaat 5640
gggttctatt gactgacaag cactgctcag gagttgggaa agagtctagc ataagcacga 5700
tagcctggag actctagtga ggtctagtct tacagacagc aaaaatcacc aggttacaaa 5760
ctacattcat ttccagtttt ctgatcaggc acaggtatga atcccttctg ttgaagagaa 5820
aagtccatgt gtttaaaata tctggtttct ccagtgctat tagcgagaag acttgagccc 5880
tatacaactc ccacctggag tgacatcctg tcttcatggt atattacata cctagacacg 5940
ctcatctcac agacttagga ctttgtcttc tgatctccat ttctgatccc acttccacct 6000
ttgccttgat agtgtcattt tcttcactgc cttggtgaca accatgttat cctctgtgta 6060
tttgagtgtt accattttca gattttacct gtatgcaaga tcacacagtc tttgtctttc 6120
tgtctggatg catgctaatc tctacacaac aacccttccc cgtcactcag atcttcctcc 6180
attaacacat acatggtgct gaagaggcta gggagcttcc cttcagtggg gagctagctg 6240
gctattgggc ctttttgact gtccaggaag gcccccaatt gctgagacaa gaacttagat 6300
tcttcattat tgactctaac tcatgtatca agcagaagct aatgaatagt tatcaacagg 6360
atcagaggtt ccagtgtaag acactttgac atgaaagaac ggaggaagga cagatggatg 6420
cataaaagca ggaccactgc cccaggaagg tcctggaaac tgatgcaggg caaaggacag 6480
gttataaacc aaatcttagg gagtcaggaa gagcacagag gagctcaacc aactgaccac 6540
tgcttagggg ctaccaaccc aatcctccct gtgggaacag ctaagctatc agccaagggt 6600
aataaacagg caggacctgt ggatgacatg gagagcatag ggaccctggg tccagccttt 6660
agcacctgca ctctcaggat actccaccat tgtgtcttag agagcctagg gatactgggt 6720
ccagcctttg gtaccttcac tctcagggta ccccatcact gtgtcttgga gagcctaggc 6780
accctgggtc cagccttcag tacctgcgct ctcaggacac cccaccattg tctcttgccc 6840
cgtctcttct tcctcttcct ccctttcatt gtctcttctc tgtttctttc ttgactctcc 6900
tttcccctca caccctcact ctagttctcc ccttccctct ctgcatcacc ctattctctc 6960
tgtggtccct ccactttcct ttatctctca tgcttctctc ctccctcaaa tacttgtcac 7020
ccactatact tcaggggcca gctctagtga caaagctgtt aatagcaaga ctctcagatc 7080
tccaacggct cagaggagcc agacccacca agaactctct ccaggtccaa tttcaggttc 7140
cttcgaaagc tttcagcaaa tgctcaggga acatgccact aacaagaaga tgcaaattcc 7200
agttgagagt gggaaaggcc cttgcgtagg tcccatcttc caggccaagg tcagaggggc 7260
tctgtgtaat ccggattgac agggctcaga acaatgtttt gtttttaagg tttatttatt 7320
ttaggtgtta gtgtctttgc ttgcatgacc ttatgtgcat catgtgtgtg caggttcctg 7380
atgacagtag aggagggctt tgaatccctg gggataggaa gttacaggaa attataagct 7440
gctttgtggg tcttctagct ttcccaacag aagtgaatgc tcttcaccac tgagccatct 7500
ctctaggccc aagagacatt gctttatgga tataattgtg tgtgtgtgtc aacattgagg 7560
aaagggaaat aaaaaaaaaa cttcagccgc taaggttgta cagtttcact aattgctact 7620
tttagttgtg ataaaatggc aggtgcttca acatttatat atacaaaaac ttccctgctg 7680
gtggttcaac tgtgagaact ggggtaagtg ggtgagttct ctttttctgt ctctgtctct 7740
gtctctctcc ttccattctt tcttaaagga aataaacatt gcagctgggt tatagctcat 7800

caatatggaa gttacagaag tgaaaaaagg cattgccttg gtgggtggtg ttaccagctg 7860
atttttggtt gtcctgcaag gaggtctggg gactggctgc tctgtctctg tctgtatgag 7920
tgagggaagt ctggggagca gattccctaa ccttcagcct ggcctggttc ctgagtgaac 7980
ccagcctctc tggtcctagt agctttttcc aaacaggaat ctgagtggtg acagggaaca 8040
agtaccagcc cattgcttaa gtgccagggt tagtgagggc aggaagctgc catagctggg 8100
attagtagtt gtattggatg taggaagtcc tatcctggga cagctaatcc ttaatgcttc 8160
actggagatt ttcaatgaga aatttatccc acggcccata tggccccatc cttttgtctc 8220
caacagccaa gtattttcca ttagaggaga cttcctgtac acttgatgga tgctcattcc 8280
aaggtgactt ggggcagtca gtacagactt gggatgacct ctgacagcct aacctctccc 8340
caacaagggc cctctatgtt tgctatgtaa tgtaatgtca gacattgtca ggagtgtccg 8400
cagcacagcc tgcccagtgt gagggctctc ataggtttcc cactgtctta tctacacagg 8460
gataacgagg aggtaagctg cagttcccag tctcacttca cagaggaaga gataacccca 8520
tcccaggtca tgtagccagc agtggaaaga atgaggattt gaactcaggt cttccaagtc 8580
ccattgatag catctcctca caagtccctt gccaccctca cgatgcctta gacacttgcc 8640
tgccctttat actaaggaga tgcaggtaca aggggtttac ccatgtagca gctgaggcag 8700
ctggggatag ataccagcag caggcctgat gtcaccactc taactccagc atccccagtc 8760
tgtgttcctg gagtgtgaaa atccctactt aacaagattg tgcaacagtc cttggctctg 8820
tgacccatag ctggaaacag gattctcatt gatttgtgga acatggtggc agccagccaa 8880
aaagagggtc tgcatacaga agacacgtgt ggcaaggcca cagcagactc tgactacctt 8940
agcttacaga attacaaggt cataatgtcc tctgctttgg tcacctcatg ttaaggacag 9000
gccctaatga agatggggca gaagactgaa ggaatggcca accaataact ggcccaactt 9060
gagacccatc ctacaggcaa gcatcaattc ctgacactac taatgatact ctgttatgct 9120
tgcagacaga agcctagcat aactatcctc cgagaggtcc acccagcaac tgactgaaac 9180
agaaaaagat atccacaggc aaacagtgga tggaggtcag ggactattat gggagagctg 9240
tgggaaggat taaaaaccct gaaggggata ggaaccccac aggaagacca acagagtcaa 9300
ctaagagacc tgtgggagct ctcagagact gagccaccaa ccaaagagca tacacaggcc 9360
ggtccgaggc acctggcacg tgtgaagcag acatgcagct cagtctccat gtaggtcctc 9420
caataagcgg tagcctgact gcagtatcca atccccaaca gggctgcata gtctggcctc 9480
agtgggggag gatgccccta atcctgcaga gacttgatga gtggagagct atccaggggg 9540
aacccaccct ctctgagaag ggaatgggga tgggggaggg actctgtgaa gaggggacaa 9600
ggacaaacaa gaacctcaaa taggtcaggc cctaaaggct tgctaagtag cagtggccca 9660
gctctgtcct gttcctcagc ccaaggctca gctcccacct gtttctgtgt ttttctggct 9720
tttcatgggc ctaggacttg gtgaccagtt caaacaatgg ggcctgtgga agacacaata 9780
tacaagacta gggacattcc tgttctgctg actatccata gcctgatgta ggtggaagga 9840
cccaatcact ggatttctac ccttgcacaa ccttgacagc tgagggcctc tcagaaacct 9900
atttcttcca ctgaaaaatg agactctcaa atgaacgtcg tgacaatcat caggcttatt 9960
aaagaggtgt atctaacctg aatggcaagc agacagcagg caaatgtctg tatcaacctc 10020
taggaaggac aagaactgct cactgctgcc ccccaggagg ccatttgctg aaacagctgc 10080
tctcctgctg gtgcacaggc cctgccttct cattgcagcc acagcccctt cctgtctgaa 10140
cctcctgtca ggtcactggg aaacagatca agatggaaca ggacagctcc tgatggtaaa 10200
taaaaaacag tggtcatggc tattcatagg ggtttatgct tcttcagtcc acactgtgaa 10260
gagctgtggg catgaaccac agtgttcgag gtagagttgg ggttctgaaa ttcacagtgg 10320
ggtgagctca gtaaatgtga gctggaggtc actcgtgaga cacacagtcc tgctgcttct 10380
gttcccaata tcctgaggag acgacacatc tactttgttc agaggccaca gtctagttga 10440
cctgagagtt accagtttct tatttgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg 10500
tgttgttcgt gtgtgagtgc aggtgcacat atgatagcgt acacgttgag gtcagaggat 10560
aactatcagg cgttgtcccc tcctactttt cctcggactc tggagaacaa acatgggtcc 10620
ttattccagg ggagcaagtc gctgttggct gacacatctt gctcacatac attttaccta 10680

gacaatggag cctccatcag agtattactt tagctcctca ccgatggcaa tgcaccacct 10740
ctctacccac ataggagttg ggtctccaca cacccccaca ccccccttcac caaaacgttt 10800
tcagttactt tatctggtaa agttcatcag agaatgaagc cagtattaag aacatggaat 10860
catttgggaa cctggatcta gcaatacccc accctagatg gagttgctga gttttcacct 10920
cagattataa ttccccccta gcttctatgg tttattctga aaccagggga actcgattcc 10980
tccctttgga ccacagacat cctggcttgt gaattcacat gtcatctact gctaatccat 11040
tggtagtatg tggctcacag agacacacta cagtcatggc caatgtcaag gtaggacaga 11100
tgtgaatcat tcccccagtc ctgctgtttt catgactaac cctcctcagc acagtgacca 11160
tgaacctact tttcccctcc ttttatttttt agaattgctg gaattttcta ttttgagaaa 11220
taatagcctt gggcagcatt aaacaaaatc atctagaaag ctggtttaaa atacagatgg 11280
ttgagtcagt gaaagagtga ggaatgtcat tattggcccc tcacagaggc tggctcactc 11340
cagcagaggt ggttgaagct cttggacacg ggtcaggtgc ataggaaagg tngtctggga 11400
cactgagaac cacaattgaa caaacagaac tgttggcttt ttttttttta aatgagttct 11460
caaaaaatga ctggctagct taggcaaata cttcgagcca acccaacaga acattcttcc 11520
attgattcat tctggatctt ctttctagac aatactgaac tgacccctg ttggcagtct 11580
caagtttgac aacatagggc tttgaacttg gcacaaggtc catcactgtc acccaagcat 11640
cctgggtgac ctttgggttg gaatatcttg gctaacctta gatattttct ttggagtatc 11700
tttagaacat ccaggaaata gggcttgatt ctcatcctgg gaccacaata taagtcaccc 11760
tagaatccca ggagatcgtg cagagaaaca aggatctctc tcgtgtgcat ccttcttcaa 11820
agcagtgagt agtgactcca ctaaactgag ttcccatctg agagtccaca ggaggctttg 11880
gggcaagaag cagagggaag gcactgtttg tgttggtaaa gttttgactc taacaaattt 11940
gaagacatag atgacattgt gtcagactaa caacaaccta gactcatgtg ggttctgttt 12000
agggatcaga ttttattcat caatgacttg tcttagtgta tagagaaagg cttcctactg 12060
gagtgtaggc tcaataatga cagaagagat agctatttcc cctagggact gtgctgctcc 12120
aagtttggtg gagaaaggca gtggggaacc tagatgtgct ctctggggag ggggtctgaa 12180
gctggcttca tagaaggtgt gaagttttgc tgaaacatct aaacagaatt atagcttagg 12240
aaagtgagca ggcaaggcag ggaatgtgtt gcatatgtat atgtacatga atatattatg 12300
ttatagatac acacacattt gaacctcatt tgcagatgac agaaaatagg ttattttgcc 12360
tctcttaact gctaagcaca atgacttcca gttccatcca tttcctgaaa tgccacaatt 12420
tcatttttca ttgtggctga ataaaattcc attgcagact gggccctact tcatccactc 12480
ctgagggcag gcatatcccc tggctccatt tcttacctat tgtgaagaga agtgcaactg 12540
tcttgttgaa aggcaagcgt gagagaggca ggcactaatt gtgggttttt gtttcttctt 12600
cctgctatga ctctccattt gtcagaacca aagatcgata aaagccgcca ccatgaaagc 12660
catcttaatc ccatttttat ctcttctgat tccgttaacc ccgcaatctg cattcgctca 12720
gagtgagccg gagctgaagc tggaaagtgt ggtgattgtc agtcgtcatg gtgtgcgtgc 12780
tccaaccaag gccacgcaac tgatgcagga tgtcacccca gacgcatggc caacctggcc 12840
ggtaaaactg ggttggctga caccgcgcgg tggtgagcta atcgcctatc tcggacatta 12900
ccaacgccag cgtctggtag ccgacggatt gctggcgaaa aagggctgcc cgcagtctgg 12960
tcaggtcgcg attattgctg atgtcgacga gcgtacccgt aaaacaggcg aagccttcgc 13020
cgccgggctg gcacctgact gtgcaataac cgtacatacc caggcagata cgtccagtcc 13080
cgatccgtta tttaatcctc taaaaactgg cgtttgccaa ctggataacg cgaacgtgac 13140
tgacgcgatc ctcagcaggg caggagggtc aattgctgac tttaccgggc atcggcaaac 13200
ggcgtttcgc gaactggaac gggtgcttaa ttttccgcaa tcaaacttgt gccttaaacg 13260
tgagaaacag gacgaaagct gttcattaac gcaggcatta ccatcggaac tcaaggtgag 13320
cgccgacaat gtctcattaa ccggtgcggt aagcctcgca tcaatgctga cggagatatt 13380
tctcctgcaa caagcacagg gaatgccgga gccggggtgg ggaaggatca ccgattcaca 13440
ccagtggaac accttgctaa gtttgcataa cgcgcaattt tatttgctac aacgcacgcc 13500
agaggttgcc cgcagccgcg ccaccccgtt attagatttg atcaagacag cgttgacgcc 13560

ccatccaccg caaaaacagg cgtatggtgt gacattaccc acttcagtgc tgtttatcgc 13620
cggacacgat actaatctgg caaatctcgg cggcgcactg gagctcaact ggacgcttcc 13680
cggtcagccg gataacacgc cgccaggtgg tgaactggtg tttgaacgct ggcgtcggct 13740
aagcgataac agccagtgga ttcaggtttc gctggtcttc cagactttac agcagatgcg 13800
tgataaaacg ccgctgtcat taaatacgcc gcccggagag gtgaaactga ccctggcagg 13860
atgtgaagag cgaaatgcgc agggcatgtg ttcgttggca ggttttacgc aaatcgtgaa 13920
tgaagcacgc atacccgctt gcagtttgta aggtacccgg ggatcacaac ttgccctctg 13980
aagaggaaga acagaaggat gccacaactc tcctgctggc tactctccag tggtttcatc 14040
ttacttctga tggcatttcc ctctagaaag tgctactatc atccacacat ttctacctga 14100
gaccacccaa aggaccctcc caaattctct tcctctctga gtagtctcca cacctgttac 14160
caccatccca gaattaaaat cctaactgca ctctggcgtg tgacttgcct cagtccttgc 14220
aataagagtt gttggcagtg ccaggcgtgg tggcgcacgc ctttaattcc agcacttggg 14280
aggcagaggc aggcggattt ctgagttcga ggccagcctg gtctacagag tgagttccag 14340
gacagccagg gctatacaga gaaaccctgt gtcgaaaaac caaaaaaaaa aaaaaaagtt 14400
gttggcagag tgtgggttat ataccaggtg gagatttcaa atgagtggct gaagctgtag 14460
ccagaaggaa cttagaggat agctcataac ttaaaaagaa atgtagagag tagcagaaac 14520
attgagagag tgggcacaca gccactgtgt gaatgtggca gaacacaatc cagccagcta 14580
tacatgcata agtgtatatt ggcgccatcc tgactgatga gacacaggaa aacagataga 14640
cggggttagg tggccatggc ctttcctgcc tgcctcttcc taagggtcat ctcaagacct 14700
tatgctctct taactcttcc attgctactt agcttctaga tatcacctcc agattagtct 14760
ccttgggtac atcagtgatc ctggtgatat ccagggcttc ctgattccat ctttgtcata 14820
gaggctgcaa ctaaagaggt cttcttaata cttcacaccc tgatgccaaa aggaagacac 14880
agaagttcac agaggtgaag tgattcatgt aggacataca gtgagcaagc atcagggtcc 14940
ggattatctg actctactct aacttttatg taaatgtgct ttatgccatt aacactgtca 15000
ttcctgtgct tcagctctgg gagactccca agcactctta ggcacaagcc acaattaagg 15060
gactctgaca ctctgcattg attaattagc atggtggtct ctatgtttcc agattcatga 15120
ttgtttcact ttccatatag gctatgaagg gtgtgaggaa atttttgggg gacagaattg 15180
gaggcaatcc acctctctca ggaagcctct atctggaaaa gcttacaact cagggacagt 15240
aactgtaggc ccagtccttg gtgtccaaaa tgggttttat ggtttgaatc tgcaaagcct 15300
tccatgtgct caaaggtttg aacatggagc ctcctcctgg taacactgta ttggaggctt 15360
ttgagactgg atgctctttg gtcccatgtt ttgctacatc atctgtcaag atatgaccca 15420
ggcatgctac cagctaccac agactatgcc tctccagctt tcatgttctc cccaccatga 15480
tagacttgta tctcctaaaa atggaatcaa agcaaacttt tcctgcatta agttttttt 15540
tttctgttaa gtgtttggtc acagggacaa gaaaacactc aatacagata attagtacca 15600
gagttgaggt tcattgctct agcaagttgg atcaaatttt tagggctttg gaactgattt 15660
ataagagaca tgtagaagag tctgaagctg tgggctacag aagtgtcacc agtttttaag 15720
aatagtttaa tacaccatgg gaattgtgaa aatcagaatg ctcacacaaa ggcagacagg 15780
aaaacgtgag catgtggcgt gtgagagggc ataagaagga acctaggggg aaatgagcta 15840
gaagccattc ggctacgtta gggaacgtgt gtggctgtgc ttggcccatg ccctggcaat 15900
ctgaatgagg ccaaatttta aaggagtgga ctaactcgat tgtcagagaa aatatcaaga 15960
cagaccacca ctcaggctat gccgtgtttg tgaccgacca gctactctta gccagctcta 16020
ttgtgaaatt ccagagcaat tatcagagca tgaagataca tacagtttag tgaagtaagg 16080
ggtgtgggtc cctaagtgga tggtgcataa atctatgtag gtgatgccta agtgacactt 16140
gataatccaa aatatcagca atgtggaatg tcttccaagg agacctgtag acacacattt 16200
tagaactttg ctcatggctg taataaatag ctagctagaa atcatttcct gaagaggtta 16260
gtctgagtta cggttccagg gcaaacattc agtgatggca aggaaggcat tgcagtcagg 16320
agccaaaggt cagctggtca cattgcatca agagtagaga gtcagagtgt gagtagaaag 16380
aggatacagg ttataaaacc tcactgtcca ctctcagcaa tccatttttct cctaaaaggc 16440

tttaccttct aaagatttta gtcttcaaaa ccagtaccag tagcctggga acaaaagttg 16500
aaacaaatga gcctttgtgg ggcatttcac acttaaaaca gggcatcacc taggaggagc 16560
cctgtgtgca gtaggaagtg tggcctctgt gtcaggaatg ctcaggctaa taaggggtcc 16620
tctatctgag ggaccctatg aagattcaac aagtagttgt gagaattccc tgtaaatgga 16680
tgctaccaat ttgacatttg tagacctgct attgtgtgct tctttattgg gctctcccat 16740
ctcccaactt tccaacccat attccacatt aatcccttcc accaccatgc aacactaggt 16800
aggagagaag gaaggttaga agagaaagtg ggtatagatc tatttagact acttcctgct 16860
gattaggggc aagtccaatc gtcattgtca ggatacctcc aaccagcaac cagcaaacca 16920
gcaaatcaga aacagcaaaa gcagccaaca aggcagcact aaccagcagg attggggtcg 16980
gtagcgtggg agcagtcact actggtcttc tcatggcttt ggcattaata ctctctcaag 17040
aaattccgta atttttccc caccacctga aattccgtaa ttttaaatgc aaactatcta 17100
cagctggcaa aaatcacatc tctcctagag cacaagacaa atcatagtta ctggctattt 17160
gcaatctgaa gcatctcaat atcccacacc tgggattaaa acaaaaacat attcacatca 17220
cataactgtt tttttttcc aatttttat taggtatttt ctttatttac atttcaaatg 17280
ctatcccgaa agtcccctat accctcccac ctccctgctc ccctacacac ccactcccac 17340
tttttgaccc tggagttccc cggtactggg gcatataaag tttgcaagac caaggggcct 17400
ctcttcccag tgatggccga ctaagccatc ttctgctaca tatgcagata gagacacgag 17460
ctctgggggt actagttagt tcatattgtt gttccaccta tagggtcgca gaccccttca 17520
gctccttggg tactttgtct agctcctcca ctgggggctc tgtgttttat ctaatagatg 17580
actgtgagca tccacttctg tatttgacag gcactggcct agcgtcacat gagccagcta 17640
tatcagggtc ctttcagcaa aaccttgctg gcatgtgcaa tagtgtctgc gtttggtggt 17700
tgattatggg atggatccac tagttctaga gc 17732

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/55, 9/16, 1/19, 1/20, 1/21, A01H 5/00, A23K 1/00, 1/165, C12Q 1/44, 1/68, C12N 15/82 // (C12N 1/19, C12R 1:84) (C12N 1/21, C12R 1:01, 1:125, 1:19)

A2

(11) International Publication Number: WO 97/48812

(43) International Publication Date: 24 December 1997 (24.12.97)

(21) International Application Number: PCT/CA97/00414

(22) International Filing Date: 13 June 1997 (13.06.97)

(30) Priority Data:

60/019,735	14 June 1996 (14.06.96)	US
08/744,779	6 November 1996 (06.11.96)	US
08/862,531	23 May 1997 (23.05.97)	US

(71) Applicant *(for all designated States except US)*: HER MAJESTY THE QUEEN IN RIGHT OF CANADA, represented by THE DEPARTMENT OF AGRICULTURE AND AGRI-FOOD CANADA [CA/CA]; Lethbridge Research Station, P.O. Box 3000, Main Lethbridge, Alberta T1J 4B1 (CA).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: CHENG, Kuo, Joan [CA/CA]; 2015 - 6th Avenue South, Lethbridge, Alberta T1J 1C2 (CA). SELINGER, Leonard, Brent [CA/CA]; 114 Laval Court West, Lethbridge, Alberta T1K 4G3 (CA). YANKE, Lindsey, Jay [CA/CA]; 1520 - 23rd Avenue N. #18, Lethbridge, Alberta T1H 4X9 (CA). BAE, Hee, Dong [KR/KR]; Eung-Sup Shim, 114-5, Chi-dong, Paldal-gu, Suwon, Seoul 442-060 (KR). ZHOU, Luming [CA/US]; 1227 East Brickyard Road #203, Salt Lake City, UT 84106 (US). FORSBERG, Cecil, Wallace [CA/CA]; 44 Hands Drive, Guelph, Ontario N1G 3H3 (CA).

(74) Agent: McKAY-CAREY, Mary, Jane; McKay-Carey & Company, 10155 - 102nd Street, 2125 Commerce Place, Edmonton, Alberta T5J 4G8 (CA).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

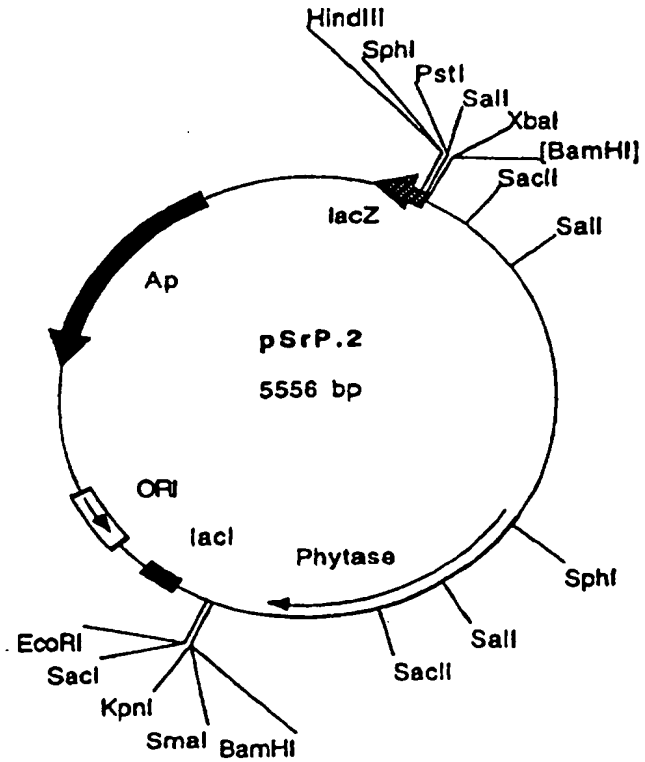
Published

Without international search report and to be republished upon receipt of that report.

(54) Title: DNA SEQUENCES ENCODING PHYTASES OF RUMINAL MICROORGANISMS

(57) Abstract

Phytases derived from ruminal microorganisms are provided. The phytases are capable of catalyzing the release of inorganic phosphorus from phytic acid. Preferred sources of phytases include *Selenomonas, Prevotella, Treponema and Megasphaera*. A purified and isolated DNA encoding a phytase of *Selenomonas ruminantium* JY35 (ATCC 55785) is provided. Recombinant expression vectors containing DNAs encoding the phytases and host cells tra nsformed with DNAs encoding the phytases are also provided. The phytases are useful in a wide range of applications involving the dephosphorylation of phytate, including, among other things, use in animal feed supplements.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

DNA SEQUENCES ENCODING PHYTASES OF RUMINAL MICROORGANISMS

Field of the Inv ntion

This invention relates to phytases derived from ruminal microorganisms.

Background of the Invention

Although the plant constituents of livestock feedstuffs are rich in phosphorus, inorganic phosphorus supplementation is required to obtain good growth performance of monogastric animals. Phytic acid (*myo*-inositol hexaphosphoric acid) generally occurs as a complex of calcium, magnesium and potassium salts and/or proteins, and is the predominant form of phosphorus in cereals, oil seeds, and legumes, and accounts for 1 to 3% of the seed dry weight and 60 to 90% of the total phosphorus present in seeds (Graf, 1986). However, monogastric animals (e.g., swine, poultry and fish) utilize phytate poorly or not at all because they are deficient in gastrointestinal tract enzymes capable of hydrolyzing phytate. Phytate passes largely intact through the upper gastrointestinal tract, where it may decrease the bioavailability of nutrients by chelating minerals (e.g., calcium and zinc), binding amino acids and proteins (Graf, 1986) and inhibiting enzymes. Phytate phosphorus in manure poses a serious pollution problem, contributing to eutrophication of surface waters in areas of the world where monogastric livestock production is intensive.

Production inefficiencies and phosphorus pollution caused by phytate may be effectively addressed by phytase supplementation of diets for monogastric animals. Phytases catalyze the hydrolysis of phytate to *myo*-inositol and inorganic phosphate, which are then absorbed in the small intestine. In addition to decreasing phosphorus supplementation requirements and reducing the amount of phytate pollutants released, phytases also diminish the antinutritional effects of phytate.

Phytases are produced in animal and plant (predominantly seeds) tissues and by a variety of microorganisms (U.S. Patent No. 3,297,548; Shieh and Ware, 1968; Ware and Shieh, 1967). Despite the array of potential phytase sources, only soil fungi (*Aspergillus niger* or *Aspergillus ficuum*) are currently used for commercial production of phytase. The phytase produced by *A. ficuum* possesses greater specific activity (100 units/mg of protein (wherein units are defined as µmoles of

phosphate released per minute)) and thermostability compared to those phytases that have been characterized from other microorganisms (European Patent Application No. 0,420,358 (van Gorcum *et al.*, 1991) and U.S. Patent No. 5,436,156 (van Gorcum *et al.*, issued July 25, 1995)). The *A. ficuum* phytase is an acid phytase and exhibits little activity above pH 5.5 (Howson and Davis, 1983; van Gorcum *et al.*, 1991). Consequently, activity is limited to a relatively small region of the monogastric digestive tract, in which the pH ranges from 2-3 (in the stomach) to 4-7 (in the small intestine).

Although the idea of phytase supplementation of monogastric diets was proposed more than 25 years ago (U.S. Patent No. 3,297,548, Ware and Shieh, 1967), the high cost of enzyme production has restricted the use of phytase in the livestock industry. In North America, supplemental phytase is generally more expensive than phosphorus supplements. In some circumstances, the cost of phytase utilization may be partially offset if the use of this enzyme also decreases the need for supplementation of a second nutrient such as calcium. The use of phytase in North America is likely to increase as swine and poultry populations increase and as public pressures force a reduction in pollution associated with livestock production. Higher costs of phosphorus supplements and legislation requiring the use of phytase have made the use of this supplement more common in Europe and parts of the Orient than in North America. Governments of the Netherlands, Germany, Korea and Taiwan have enacted or are enacting legislation to reduce the phosphorus pollution created by monogastric livestock production.

A more effective means of increasing phytase utilization is through cost reduction. The cost of phytase can be reduced by decreasing production costs and/or producing an enzyme with superior activity. Recent advances in biotechnology may revolutionize the commercial enzyme industry by offering alternative, cost effective methods of enzyme production. Application of recombinant DNA technology has enabled manufacturers to increase the yields and efficiency of enzyme production, and to create new products. The original source organism need no longer limit the production of commercial enzymes. Genes encoding superior enzymes can be transferred from organisms such as anaerobic bacteria and fungi, typically impractical for commercial production, into well characterized industrial

microbial production hosts (e.g., *Aspergillus* and *Bacillus* spp.). As well, these genes may be transferred to novel plant and animal expression systems.

Unlike monogastric animals, ruminants (e.g., cattle, sheep) readily utilize the phosphorus in phytic acid. It has been demonstrated that phytases are present in the rumen, and it has been proposed that ruminants reared on high grain diets (rich in phytate) do not require dietary phosphorus supplementation due to these ruminal phytases. A single report has attributed this phytase production to ruminal microorganisms (Raun *et al.*, 1956), but overall, the unique capacity of ruminants to utilize phytate has largely been ignored. Raun *et al.* (1956) prepared microbial suspensions by centrifugal sedimentation (Cheng *et al.*, 1955). Those microbial suspensions were almost certainly contaminated with microscopic particles of plant material. Since plants produce phytases, the study was inconclusive as to whether plant phytases or microbial phytases produced the observed activity. Although Raun *et al.* have raised the possibility that ruminal phytase production may be attributable to ruminal microorganisms, this possibility has not been explored.

In view of the foregoing, there remains a need for low cost phytases having biochemical characteristics well suited for use in animal feed supplements.

Summary of the Invention

The inventors have discovered that the rumen is a rich source of microorganisms which produce phytases having biochemical characteristics (such as temperature and pH stability, low metal ion sensitivity and high specific activity) desirable for industrial applications such as animal feed supplementation and inositol production. Ruminal microorganisms tolerate anaerobic conditions and may be either facultative or obligate anaerobes. Ruminal microorganisms may be prokaryotes (i.e. bacteria) or eukaryotes (i.e. fungi, protozoa). As used herein, the term "ruminal microorganisms" includes microorganisms isolated from the digesta or feces of a ruminant animal.

Ruminal bacterial species which have been identified as providing particularly active phytases includes *Selenomonas ruminantium, Prevotella sp, Treponema bryantii and Megaphaera elsdenii. Prevotella* and *Selenomonas* are Gram negative anaerobic rods from the family Bacteriodaceae.

In accordance with the present invention, DNA sequences encoding novel and useful phytases derived from ruminal microorganisms are provided.

A phytase gene (*phyA*) from *Selenomonas ruminantium* strain JY35 has been cloned and sequenced, and the nucleotide sequence of the *phyA* gene is provided. The invention extends to DNA sequences which encode phytases and which are capable of hybridizing under stringent conditions with the *phy*A gene sequence. As used herein, "capable of hybridizing under stringent conditions" means annealing to a subject nucleotide sequence, or its complementary strand, under standard conditions (ie. high temperature and/or low salt content) which tend to disfavor annealing of unrelated sequences. As used herein, "conditions of low stringency" means hybridization and wash conditions of 40 - 50°C, 6 X SSC and 0.1% SDS (indicating about 50 - 80% homology). As used herein, "conditions of medium stringency" means hybridization and wash conditions of 50 - 65°C, 1 X SSC and 0.1% SDS (indicating about 80 - 95% homology). As used herein, "conditions of high stringency" means hybridization and wash conditions of 65 - 68°C, 0.1 X SSC and 0.1% SDS (indicating about 95-100% homology).

As used herein, the term "phytase" means an enzyme capable of catalyzing the removal of inorganic phosphorus from a *myo*-inositol phosphate.

As used herein, the term "*myo*-inositol phosphate" includes, without limitation, *myo*-inositol hexaphosphate, *myo*-inositol pentaphosphate, *myo*-inositol tetraphosphate, *myo*-inositol triphosphate, *myo*-inositol diphosphate and *myo*-inositol monophosphate.

As used herein, "phytate" means the salt of *myo*-inositol hexaphosphoric acid.

The invention extends to the *S. ruminantium* JY35 (ATCC 55785) organism itself, and to methods for identifying and isolating this and other ruminal microorganisms exhibiting phytase activity as well as methods for isolating, cloning and expressing phytase genes from ruminal microorganisms exhibiting phytase activity using part or all of the *phy*A gene sequence as a probe.

The invention further extends to methods for assaying phytase production by a microorganism whereby false positive results caused by microbial acid production are eliminated. Colonies of microorganisms are grown on a growth medium containing phytate. The medium is contacted with an aqueous solution of cobalt

chloride and the medium is then examined for zones of clearing. Preferably, rather than examining the medium immediately, the solution of cobalt chloride is removed and the medium is contacted with aqueous solutions of ammonium molybdate and ammonium vanadate and then examined for zones of clearing. False positive results which occur when acid-forming microbes produce zones of clearing are avoided.

The invention extends to expression constructs constituting a DNA encoding a phytase of the present invention operably linked to control sequences capable of directing expression of the phytase in a suitable host cell.

The invention further extends to host cells which have been transformed with, and express, DNA encoding a phytase of the present invention, and to methods of producing such transformed host cells. As used herein "host cell" includes animal, plant, yeast, fungal, protozoan and prokaryotic host cells.

The invention further extends to transgenic plants which have been transformed with a DNA encoding a phytase of the present invention so that the transformed plant is capable of expressing the phytase and to methods of producing such transformed plants. As used herein, "transgenic plant" includes transgenic plants, tissues and cells.

Phytases of the present invention are useful in a wide variety of applications involving the dephosphorylation of phytate. Such applications include use in animal feed supplements, feedstuff conditioning, human nutrition, and the production of inositol from phytic acid. Phytases of the present invention may also be used to minimize the adverse effects of phytate metal chelation. The high phytate content of certain feedstuffs such as soy meal decreases their value as protein sources for fish, monogastric animals, young ruminants and infants because the phytate decreases the bioavailability of nutrients by chelating minerals, and binding amino acids and proteins. Treatment of such feedstuffs with the phytases of the present invention will reduce their phytate content by phytase mediated dephosphorylation, rendering the feedstuffs more suitable for use as protein sources. Accordingly, the invention extends to novel feed compositions comprising feedstuffs treated with a phytase of the present invention, and feed additives containing a phytase of the present invention. Such feed compositions and additives may also contain other enzymes, such as, proteases, cellulase, xylanases and acid phosphatases. The

phytase may be added directly to an untreated, pelletized, or otherwise processed feedstuff, or it may be provided separately from the feedstuff in, for instance, a mineral block, a pill, a gel formulation, a liquid formulation, or in drinking water. The invention extends to feed inoculant preparations comprising lyophilized microorganisms which express phytases of the present invention under normal growing conditions. With respect to these feed inoculant preparations, "normal growing conditions" mean culture conditions prior to harvesting and lyophilization of the microorganisms. The microorganisms express phytases during growth of the microbial cultures in large-scale fermenters. The activity of phytases in the microorganisms is preserved by lyophilization of the harvested microbial concentrates containing the phytase.

The invention further extends to a method for improving an animal's utilization of dietary phosphate by feeding the animal an effective amount of a phytase of the present invention. As used herein "an effective amount" of a phytase means an amount which results in a statistically significant improvement in phosphorus utilization by the animal. Phytate phosphorus utilization may be evidenced by, for instance, improved animal growth and reduced levels of phytate in animal manure.

Brief Description of Drawings

Figure 1 is a photograph showing the effect of counterstaining agar medium containing phytate on zones of clearing produced by acid production or phytase activity. Phytate agar was inoculated with *S. bovis* (top of left petri dish) and *S. ruminantium* JY35 (bottom of left petri dish) and incubated for 5 d at 37°C. The colonies were scraped off and the medium counterstained with cobalt chloride and ammonium molybdate/ammonium vanadate solutions (right petri plate).

Figure 2 is a graph illustrating the growth (protein) and phytase production of *S. ruminantium* JY35 in modified Scott and Dehority (1965) broth.

Figure 3A, 3B and C show transmission electron micrographs of cells from a mid-exponential phase culture of *S. ruminantium* JY35 incubated for reaction product deposition by phytase using sodium phytate as the substrate. Untreated control cells are shown for comparison in Figures 3D, 3E and 3F.

Figure 4 is a graph illustrating the phytase pH profile for washed *S. ruminantium* JY35 cells in five different buffers.

Figure 5 is a graph illustrating the pH profile of *S. ruminantium* JY35 $MgCl_2$ cell extract in five different buffers.

Figure 6 is a graph illustrating the temperature profile of *S. ruminantium* JY35 $MgCl_2$ cell extract.

Figure 7 is a graph illustrating the effect of ions (10 mM) on *S. ruminantium* JY35 phytase activity (Ctr = control).

Figure 8 is a graph illustrating the effect of sodium phytate concentration on *S. ruminantium* JY35 phytase activity.

Figure 9 is a zymogram developed for confirmation of phytase activity. Concentrates (10 x) of *S. ruminantium* JY35 $MgCl_2$ extract (lanes B - E), low molecular weight markers (lane F, BioRad Laboratories Canada Ltd, Mississauga, Ontario) and *A. ficuum* phytase (Sigma, 1.6 U, lane A) were resolved by SDS-PAGE in a 10% polyacrylamide gel. Lanes A to E were stained for phytase activity and Lane F was stained with Coomassie brilliant blue.

Figure 10 is a photograph of a phytate hydrolysis plate assay for phytase activities of *E. coli* DH5α transformed with pSrP.2 (top), pSrP.2Δ*Sph*I (bottom left), and pSrPf6 (bottom right). Zones of clearing were visible after incubating the plates at 37°C for 48 h.

Figure 11 is a Southern blot analysis using the 2.7-kb fragment from pSrP.2 as a probe against *Sph*1 digested pSrP.2 DNA (lane B) and *Hin*dIII digested genomic DNA isolated from *S. ruminantium* JY35 (lane C). Digoxigenin labelled *Hin*dIII digested Lambda DNA was run as a molecular weight standard in lane A.

Figure 12 is a physical map of pSrP.2. A 2.7-kb fragment, from a *Sau*3A partial digest of *S. ruminatium* JY35 genomic DNA, was cloned into the *Bam*HI site of pUC18. This fragment contains the entire gene encoding the phytase from *S. ruminatium* JY35. The location of a *Bam*HI site lost as a result of the ligation is indicated in square brackets.

Figure 13 is a schematic representation of the deletion analysis of the *S. ruminatium* phytase gene. The position of *phyA* is indicated by the horizontal arrow.

The hatched boxes indicate segments of the 2.7-kb *Sau*3A fragment carried by different plasmid derivatives. Phytase activity is indicated in the panel to the right.

Figure 14 is a zymogram developed for phytase activity. *E. coli* DH5α (pSrP.2) cells (lane A), *E. coli* DH5α (pSrP.2Δ*Sph*I) cells (lane B), and low molecular weight markers (lane C, BioRad Laboratories) were resolved by SDS-PAGE in a 10% polyacrylamide gel. Lanes A and B were stained for phytase activity and Lane C was stained with Coomassie brilliant blue.

Figure 15 is the nucleotide sequence of the *S. ruminantium* JY35 phytase gene (*phyA*) (SEQ ID NO. 1) and its deduced amino acid sequence (SEQ ID NO. 2). Nucleotide 1 corresponds to nt 1232 of the 2.7-kb insert of pSrP.2. The putative ribosome binding site is underlined and shown above the sequence as R.B.S. The signal peptidase cleavage site, predicted by the method of von Heijne (1986) is indicated by the ↑. The N-terminal amino acid sequence of the phytase secreted by *E. coli* (pSrPf6) is underlined.

Detailed Description of the Preferred Embodiment

The rumen is a complex ecosystem inhabited by more than 300 species of bacteria, fungi and protozoa. Screening these organisms for phytase activity requires the ability to discriminate the phytase activity of individual isolates. This may be accomplished through the assessment of pure cultures from a stock culture collection or separation and cultivation of individual cells through cultural techniques (e.g., streak plate, dilution and micromanipulation). Standard aseptic, anaerobic techniques described for bacteria, fungi and protozoa may be used to accomplish this goal.

Suitable enzyme assays are necessary for screening microbial isolates in ruminal fluid samples and from culture collections, and for cloning phytase genes. Assays for measuring phytase activity in solutions have been described in the literature. Sample solutions are typically assayed for phytase activity by measuring the release of inorganic phosphorus (P_i) from phytic acid (Raun *et al.*, 1956; van Hartingsveldt *et al.*, 1993). Phytase activity may also be detected on solid media. Microorganisms expressing phytase produce zones of clearing on agar media containing sodium or calcium phytate (Shieh and Ware, 1968; Howson and Davis,

1983). However, the solid media assays described in the literature were found to be unsatisfactory for screening ruminal bacteria for phytase activity because of the false positive reactions of acid-producing bacteria such as *Streptococcus bovis*. To overcome this problem, a two-step counterstaining procedure was developed in which petri dishes containing solid medium are flooded first with an aqueous cobalt chloride solution and second with an aqueous ammonium molybdate/ammonium vanadate solution. Following this treatment only clearing zones produced by enzyme activity are evident (Figure 1).

Using the above solutions and solid medium assays, 345 isolates from the Lethbridge Research Centre (Lethbridge, Alberta, Canada) culture collection were screened for phytase activity (Table 1). A total of 29 cultures with substantial phytase activity were identified, including 24 of the genus *Selenomonas* and 5 of the genus *Prevotella*. Twelve of these cultures (11 *Selenomonas* isolates and 1 *Prevotella* isolate) had phytase activities substantially higher than the other positive cultures (Table 2).

The phytase of *S. ruminantium* JY35 (deposited May 24, 1996 with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland, 20852-1776, as ATCC 55785) was selected for further examination and compared to a commercial phytase (Gist-brocades nv, Delft, The Netherlands) from *Aspergillus ficuum* NRRL 3135 (van Gorcum *et al.*, 1991 and 1995). The phytase of *S. ruminantium* JY35 (ATCC 55785) is constitutively expressed, exported from the cell and associated with the cell surface. The pH (Figure 5) and temperature (Figure 6) profiles of the *S. ruminantium* JY35 (ATCC 55785) phytase were comparable, if not more suited to industrial production, than are those of the commercial *A. ficuum* NRRL 3135 phytase. These results demonstrated the potential of ruminal and anaerobic microbes as sources of phytases with characteristics superior to phytases currently being produced by industry.

Microbial genes encoding selected enzymes can be cloned by a variety of methods. Gene libraries (genomic DNA and/or cDNA) are constructed by standard methods (Sambrook *et al.*, 1989; Ausubel *et al.*, 1990) and screened for the desired gene. The screening methodology may utilize heterologous probes, enzyme activity

or results generated during purification of the gene product, such as N-terminal and internal amino acid sequence data and antibodies.

Using the solid medium phytase assay developed to detect phytase activity produced by ruminal microbes, a *S. ruminantium* JY35 (ATCC 55785) gene library was screened for positive clones. Of 6000 colonies examined, a single colony was identified as a phytase positive clone by a large zone of clearing around the colony. This clone carried a 5.5-kb plasmid comprising a 2.7-kb *Sau*3A DNA fragment inserted into cloning vector pUC18. The newly isolated 2.7-kb *Sau*3A DNA fragment was used as a probe in Southern blot hybridizations. Under high stringency conditions, a discrete band could be detected for *S. ruminantium* isolate JY35 (ATCC 55785), but not for *Prevotella* sp. 46/5[2], *E. coli* DH5α or *A. ficuum* NRRL 3135.

Plasmid DNA isolated from the newly isolated clone and introduced into *E. coli* cells by transformation produced ampicillin-resistant, phytase-positive CFUs. Zymogram analysis of cell extracts from *E. coli* DH5α cells carrying the 2.7-kb *Sau*3A DNA fragment from *S. ruminantium* JY35 (ATCC 55785) revealed a single activity band with an estimated molecular mass of 37 kDa. Deletion and DNA sequence analyses were used to identify the gene (*phyA*) which encoded the phytase responsible for the activity observed in recombinant *E. coli* clones. The N-terminal amino acid sequence of the purified 37-kDa phytase expressed in *E. coli* cells carrying *phyA* matched the N-terminal amino acid sequence of the mature phytase predicted from the cloned *phyA* sequence. This indicated conclusively that the nucleotide sequence encoding the phytase had been isolated. The nucleotide sequence and deduced amino acid sequence are shown in Figure 15.

As with other genes, it is possible to use the characterized phytase coding sequence in a variety of expression systems for commercial enzyme production. Application of recombinant DNA technology has enabled enzyme manufacturers to increase the volume and efficiency of enzyme production, and to create new products. The original source organism need no longer limit the production of commercial enzymes. Genes encoding superior enzymes can be transferred from organisms such as anaerobic bacteria and fungi, typically impractical for commercial production, into well characterized industrial microbial production hosts (e.g.,

Aspergillus, *Pichia*, *Trichoderma*, *Bacillus* spp.). As well, these genes may be transferred to novel plant and animal expression systems.

Industrial strains of microorganisms (e.g., *Aspergillus niger, Aspergillus ficuum, Aspergillus awamori, Aspergillus oryzae, Trichoderma reesei, Mucor miehei, Kluyveromyces lactis, Pichia pastoris, Saccharomyces cerevisiae, Escherichia coli, Bacillus subtilis* or *Bacillus licheniformis*) or plant hosts (e.g., canola, soybean, corn, potato) may be used to produce phytase. All systems employ a similar approach to gene expression. An expression construct is assembled to include the protein coding sequence of interest and control sequences such as promoters, enhancers and terminators. Other sequences such as signal sequences and selectable markers may also be included. To achieve extracellular expression of phytase, the expression construct of the present invention utilizes a secretory signal sequence. The signal sequence is not included on the expression construct if cytoplasmic expression is desired. The promoter and signal sequence are functional in the host cell and provide for expression and secretion of the coding sequence product. Transcriptional terminators are included to ensure efficient transcription. Ancillary sequences enhancing expression or protein purification may also be included in the expression construct.

The protein coding sequences for phytase activity are obtained from ruminal microbial sources. This DNA may be homologous or heterologous to the expression host. Homologous DNA is herein defined as DNA originating from the same species. For example, *S. ruminantium* may be transformed with DNA from *S. ruminantium* to improve existing properties without introducing properties that did not exist previously in the species. Heterologous DNA is defined as DNA originating from a different species. For example, the *S. ruminatium phyA* may be cloned and expressed in *E. coli*.

It is well known in the biological arts that certain amino acid substitutions can be made in protein sequences without affecting the function of the protein. Generally, conservative amino acid substitutions are tolerated without affecting protein function. Similar amino acids can be those that are similar in size and/or charge properties, for example, aspartate and glutamate and isoleucine and valine are both pairs of similar amino acids. Similarity between amino acid pairs has been

assessed in the art in a number of ways. For example, Dayhoff et al. (1978) in Atlas of Protein Sequence and Structure, Volume 5, Supplement 3, Chapter 22, pages 345-352, which is incorporated by reference herein, provides frequency tables for amino acid substitutions which can be employed as a measure of amino acid similarity. Dayoff et al.'s frequency tables are based on comparisons of amino acid sequences for proteins having the same function from a variety of evolutionary different sources.

It is also well-known that often less than a full length protein has the function of the complete protein, for example, a truncated protein lacking an N-terminal, internal or a C-terminal protein often ha the biological and/or enzymatic activity of the complete natural protein. Gene truncation experiments involving *phy*A have confirmed that the truncated protein may retain the function of the intact protein. *Exherichia coli* clones expressing PhyA missing N-terminal amino acids 1-37 or 1058 (SEQ ID NO. 2) showed phytase positive phenotypes. In contrast, no phytase activity could be detected for a clone expressing PhyA missing acids 307-346 (SEQ ID NO. 2). Those of ordinary skill in the art know how to make truncated protein and proteins with internal deletions. In the present invention, the function of a truncated phytase protein or an internally deleted phytase protein can be readily tested using the assay described hereinbelow and n view of what is generally known in the art.

Substituted, internally-deleted and truncated rumina phytase derivatives which retain substantially the same enzymatic activity as a phytase specifically disclosed herein are considered equivalents of the exemplified phytase and are within the scope of the present invention, particularly where the specific activity of the substituted, internally-deleted or truncated phytase derivative is at least about 10% of the specifically exemplified phytase. The skilled artisan can readily measure the activity of a rumina phytase, truncated phytase, internally-deleted phytase or substituted phytase using the assay procedures taught herein and in view of what is generally known in the art.

This invention includes structurally variant phytases derived from a phytase of a rumina microorganisms, particularly those derived from a phytase specifically disclosed herein, that are substantially functionally equivalent to that phytase as assayed as described herein in view of what is generally known in the art.

Structurally variant, functional equivalents of the phytases of this invention include those phytase of rumina microorganisms having a contiguous amino acid sequence as in the phytase amino acid sequence disclosed herein (SEQ ID NO. 2), particularly those variant phytase which have a contiguous amino acid sequence of a phytase of a rumina microorganism that is a contiguous sequence at least about 25 amino acids in length.

The present invention also provides the starting material for the construction of phytases with properties that differ from those of the enzymes isolated herein. The genes can be readily mutated by known procedures (e.g., chemical, site directed, random polymerase chain reaction mutagenesis) thereby creating gene products with altered properties (e.g., temperature or pH optima, specific activity or substrate specificity).

Various promoters (transcriptional initiation regulatory region) may be used according to the present invention. The selection of the appropriate promoter is dependent upon the proposed expression host. Choices of promoters may include the promoter associated with the cloned protein coding sequence or promoters from heterologous sources as long as they are functional in the chosen host. Examples of heterologous promoters are the *E. coli tac* and *trc* promoters (Brosius *et al.*, 1985), *Bacillus subtilis sacB* promoter and signal sequence (Wong, 1989), *aox1* and *aox2* from *Pichia pastoris* (Ellis *et al.*, 1985), and oleosin seed specific promoter from *Brassica napus* or *Arabidopsis thaliana* (van Rooijen and Moloney, 1994). Promoter selection is also dependent upon the desired efficiency and level of peptide or protein production. Inducible promoters such *tac* and *aox1* are often employed in order to dramatically increase the level of protein expression. Overexpression of proteins may be harmful to the host cells. Consequently, host cell growth may be limited. The use of inducible promoter systems allows the host cells to be cultivated to acceptable densities prior to induction of gene expression, thereby facilitating higher product yields. If the protein coding sequence is to be integrated through a gene replacement (omega insertion) event into a target locus, then promoter selection may also be influenced by the degree of homology to the target locus promoter.

Various signal sequences may be used according to the present invention. A signal sequence which is homologous to the protein coding sequence to be expressed may be used. Alternatively, a signal sequence which has been selected or designed for improved secretion in the expression host may also be used. For example, *B. subtilis sacB* signal sequence for secretion in *B. subtilis*, the *Saccharomyces cerevisiae* α-mating factor or *P. pastoris* acid phosphatase *phol* signal sequences for *P. pastoris* secretion may be used. A signal sequence with a high degree of homology to the target locus may be required if the protein coding sequence is to be integrated through an omega insertion event. The signal sequence may be joined directly through the sequence encoding the signal peptidase cleavage site to the protein coding sequence, or through a short nucleotide bridge consisting of usually fewer than ten codons.

Elements for enhancing expression transcription (promoter activity) and translation have been identified for eukaryotic protein expression systems. For example, positioning the cauliflower mosaic virus (CaMV) promoter 1000 bp on either side of a heterologous promoter may elevate transcriptional levels by 10- to 400-fold. The expression construct should also include the appropriate translational initiation sequences. Modification of the expression construct to include the Kozak consensus sequence for proper translational initiation may increase the level of translation by 10 fold.

Elements to enhance purification of the protein may also be included in the expression construct. The product of oleosin gene fusions is a hybrid protein containing the oleosin gene joined to the gene product of interest. The fusion protein retains the lipophilic properties of oleosins and is incorporated in the oil body membranes (van Rooijen and Moloney, 1994). Association with the oil bodies may be exploited to facilitate purification of the recombinant oleosin fusion proteins (van Rooijen and Moloney, 1994).

A selection marker is usually employed, which may be part of the expression construct or separate from it (e.g., carried by the expression vector), so that the marker may integrate at a site different from the gene of interest. Transformation of the host cells with the recombinant DNA molecules of the invention is monitored through the use of selectable markers. Examples of these are markers that confer

resistance to antibiotics (e.g., *bla* confers resistance to ampicillin for *E. coli* host cells, *nptII* confers kanamycin resistance to *B. napus* cells) or that permit the host to grow on minimal medium (e.g., *HIS4* enables *P. pastoris* GS115 His$^-$ to grow in the absence of histidine). The selectable marker will have its own transcriptional and translational initiation and termination regulatory regions to allow for independent expression of the marker. Where antibiotic resistance is employed as a marker, the concentration of the antibiotic for selection will vary depending upon the antibiotic, generally ranging from 10 to 600 µg of the antibiotic/mL of medium.

The expression construct is assembled by employing known recombinant DNA techniques. Restriction enzyme digestion and ligation are the basic steps employed to join two fragments of DNA. The ends of the DNA fragment may require modification prior to ligation and this may be accomplished by filling in overhangs, deleting terminal portions of the fragment(s) with nucleases (e.g., *Exo*III), site directed mutagenesis, and adding new base pairs by the polymerase chain reaction (PCR). Polylinkers and adaptors may be employed to facilitate joining of select fragments. The expression construct is typically assembled in stages employing rounds of restriction, ligation and transformation of *E. coli*. There are numerous cloning vectors available for construction of the expression construct and the particular choice is not critical to this invention. The selection of cloning vector will be influenced by the gene transfer system selected for introduction of the expression contruct into the host cell. At the end of each stage, the resulting construct may be analyzed by restriction, DNA sequence, hybridization and PCR analyses.

The expression construct may be transformed into the host as the cloning vector construct, either linear or circular, or may be removed from the cloning vector and used as is or introduced onto a delivery vector. The delivery vector facilitates the introduction and maintenance of the expression construct in the selected host cell type. The expression construct is introduced into the host cells by employing any of a number of gene transfer systems (e.g., natural competence, chemically mediated transformation, protoplast transformation, electroporation, biolistic transformation, transfection, or conjugation). The gene transfer system selected depends upon the host cells and vector systems used.

For instance, the expression construct can be introduced into *P. pastoris* cells by protoplast transformation or electroporation. Electroporation of *P. pastoris* is easily accomplished and yields transformation efficiencies comparable to spheroplast transformation. *P. pastoris* cells are washed with sterile water and resuspended in a low conductivity solution (e.g., 1 M sorbitol solution). A high voltage shock applied to the cell suspension creates transient pores in the cell membrane through which the transforming DNA (e.g., expression construct) enters the cells. The expression construct is stably maintained by integration, through homologous recombination, into the *aox1* (alcohol oxidase) locus.

Alternatively, an expression construct, comprising the *sacB* promoter and signal sequence operably linked to the protein coding sequence, is carried on pUB110, a plasmid capable of autonomously replicating in *B. subtilis* cells. The resulting plasmid construct is introduced into *B. subtilis* cells by transformation. *Bacillus subtilis* cells develop natural competence when grown under nutrient poor conditions.

In a third example, *Brassica napus* cells are transformed by *Agrobacterium*-mediated transformation. The expression construct is inserted onto a binary vector capable of replication in *A. tumefaciens* and mobilization into plant cells. The resulting contruct is transformed into *A. tumefaciens* cells carrying an attenuated Ti or "helper plasmid". When leaf disks are infected with the recombinant *A. tumefaciens* cells, the expression construct is transferred into *B. napus* leaf cells by conjugal mobilization of the binary vector::expression construct. The expression construct integrates at random into the plant cell genome.

Host cells carrying the expression construct (i.e., transformed cells) are identified through the use of the selectable marker carried by the expression construct or vector and the presence of the gene of interest confirmed by a variety of techniques including hybridization, PCR, and antibodies.

The transformant microbial cells may be grown by a variety of techniques including batch and continuous fermentation on liquid or semi-solid media. Transformed cells are propagated under conditions optimized for maximal product-to-cost ratios. Product yields may be dramatically increased by manipulating of cultivation parameters such as temperature, pH, aeration, and media composition.

Careful manipulation and monitoring of the growth conditions for recombinant hyper-expressing *E. coli* cells may result in culture biomass and protein yields of 150 g (wet weight) of cells/L and 5 g of insoluble protein/L, respectively. Low concentrations of a protease inhibitor (e.g., phenylmethylsulfonyl fluoride or pepstatin) may be employed to reduce proteolysis of the over-expressed peptide or protein. Alternatively, protease deficient host cells may be employed to reduce or eliminate degradation of the desired protein.

After selection and screening, transformed plant cells can be regenerated into whole plants and varietal lines of transgenic plants developed and cultivated using known methods. As used herein, "transgenic plant" includes transgenic plants, plant tissues and plant cells.

Following fermentation, the microbial cells may be removed from the medium through down-stream processes such as centrifugation and filtration. If the desired product is secreted, it can be extracted from the nutrient medium. In the case of intracellular production, the cells are harvested and the product released by rupturing cells through the application of mechanical forces, ultrasound, enzymes, chemicals and/or high pressure. Production of an insoluble product, such as occurs in hyper-expressing *E. coli* systems, can be used to facilitate product purification. The product inclusions can be extracted from disrupted cells by centrifugation and contaminating proteins may be removed by washing with a buffer containing low concentrations of a denaturant (e.g., 0.5 to 6 M urea, 0.1 to 1% sodium dodecyl sulfate or 0.5 to 4.0 M guanidine-HCl). The washed inclusions may be solubilized in solutions containing 6 to 8 M urea, 1 to 2% sodium dodecyl sulfate or 4 to 6 M guanidine-HCl. Solubilized product can be renatured by slowly removing denaturing agents during dialysis.

Phytase may be extracted from harvested portions or whole plants by grinding, homogenization, and/or chemical treatment. The use of seed specific lipophilic oleosin fusions can facilitate purification by partitioning the oleosin fusion protein in the oil fraction of crushed canola seeds, away from the aqueous proteins (van Rooijen and Moloney, 1994).

If necessary, various methods for purifying the product, from microbial, fermentation and plant extracts, may be employed. These include precipitation (e.g.,

ammonium sulfate precipitation), chromatography (gel filtration, ion exchange, affinity liquid chromatography), ultrafiltration, electrophoresis, solvent-solvent extraction (e.g., acetone precipitation), combinations thereof, or the like.

All or a portion of the microbial cultures and plants may be used directly in applications requiring the action of phytase. Various formulations of the crude or purified phytase preparations may also be prepared. The enzymes can be stabilized through the addition of other proteins (e.g., gelatin, skim milk powder) and chemical agents (e.g., glycerol, polyethylene glycol, reducing agents and aldehydes). Enzyme suspensions can be concentrated (e.g., tangential flow filtration) or dried (spray and drum drying, lyophilization) and formulated as liquids, powders, granules, pills, mineral blocks and gels through known processes. Gelling agents such as gelatin, alginate, collagen, agar, pectin and carrageenan may be used.

Further, complete dephosphorylation of phytate may not be achieved by phytase alone. Phytases may not dephosphorylate the lower *myo*-inositol phosphates. For instance, an *A. ficuum* phytase described in U.S. Patent No. 5,536,156 (van Gorcum *et. al.*, issued July 25, 1995) exhibits low or no phosphatase activity against *myo*-inositol di-phosphate or *myo*-inositol mono-phosphate. Addition of another phosphatase, such as an acid phosphatase, to a feed additive of the present invention containing phytase will help dephosphorylate *myo*-inositol di-phosphate and *myo*-inositol mono-phosphate.

Formulations of the desired product may be used directly in applications requiring the action of a phytase. Liquid concentrates, powders and granules may be added directly to reaction mixtures, fermentations, steeping grains, and milling waste. The formulated phytase can be administered to animals in drinking water, in a mineral block, as a salt, or as a powdered supplement to be sprinkled into feed bunks or mixed with a ration. It may also be mixed with, sprayed on or pelleted with other feed stuffs through known processes. Alternatively, a phytase gene with a suitable promoter-enhancer sequence may be intergrated into an animal genome and selectively expressed in an organ or tissue (e.g. salivary glands, pancreas or epithelial cells) which secrete the phytase enzyme into the gastrointestinal tract, thereby eliminating the need for the addition of supplemental phytase.

In a preferred formulation, phytases of the present invention may take the form of microbial feed inoculants. Cultures of microorganisms expressing a native phytase, such as *S. ruminantium* JY35 (ATCC 55785), or recombinant microorganisms expressing a phytase encoded by a heterologous phytase gene are grown to high concentrations in fermenters and then harvested and concentrated by centrifugation. Food-grade whey and/or other cryoprotective agents are then admixed with the cell concentrate. The resulting mixture is then cryogenically frozen and freeze-dried to preserve phytase activity by standard lyophilization procedures. The freeze-dried culture may be further processed to form a finished product by such further steps as blending the culture with an inert carrier to adjust the strength of the product.

All or a portion of the microbial cultures and plants as produced by the present invention may be used in a variety of industrial processes requiring the action of a phytase. Such applications include, without limitation, the manufacture of end products such as inositol phosphate and inositol, production of feed ingredients and feed additives for non-ruminants (e.g., swine, poultry, fish, pet food), in human nutrition, and in other industries (soybean and corn processing, starch, and fermentation) that involve feedstocks containing phytate. Degradation of phytate makes inorganic phosphate and chelated metals available to animals and microorganisms. The action of phytase increases the quality, value and utility of feed ingredients and/or fermentation substrates that are high in phytate. The action of phytases can also accelerate the steeping process and separation processes involved in the wet milling of corn.

The phytase genes of the present invention can be used in heterologous hybridization and polymerase chain reaction experiments, directed to isolation of phytase encoding genes from other microorganisms. The examples herein are given by way of illustration and are in no way intended to limit the scope of the present invention. Efforts have been made to ensure the accuracy with respect to numbers used (e.g., temperature, pH, amounts) but the possibility of some experimental variance and deviations should be recognized.

Example 1

Isolation of ruminal bacteria

Ruminal fluid from a cannulated Holstein cow was collected in a sterile Whirlpak™ bag. Fluid may also be withdrawn from the rumen via an orogastric tube. Under a suitable anaerobic atmosphere (e.g., 90% CO_2 and 10% H_2), ten-fold serial dilutions of the rumen fluid were prepared and distributed over the surface of a solid growth medium (e.g., Scott and Dehority, 1965), and the plates were incubated at 39°C for 18 to 72 h. Isolated colonies were picked with a sterile loop and the cells were spread over the surface of fresh agar medium to produce isolated colonies. The cells from a single colony were confirmed by morphological examination to represent a pure culture and were cultured and stored in the Lethbridge Research Centre ("LRC") culture collection or used as a source of enzymatic activity or genetic material.

Example 2

Screening ruminal bacteria for phytase activity

A. Phytase assays

Sample solutions (culture filtrates, cell suspensions, lysates, washes or distilled water blanks) were assayed for phytase activity by incubating 150 µl of the solution with 600 µl of substrate solution [0.2% (w/v) sodium phytate in 0.1 M sodium acetate buffer, pH 5.0] for 30 min at 37°C. The reaction was stopped by adding 750 µl of 5% (w/v) trichloroacetic acid. Released orthophosphate in the reaction mixture was measured by the method of Fiske and Subbarow (1925). Freshly prepared colour reagent [750 µl of a solution containing 4 volumes of 1.5% (w/v) ammonium molybdate in a 5.5% (v/v) sulfuric acid solution and 1 volume of a 2.7% (w/v) ferrous sulfate solution] was added to the reaction mixture and the production of phosphomolybdate was measured spectrophotometrically at 700 nm. Results were compared to a standard curve prepared with inorganic phosphate. One unit ("Unit") of phytase was defined as the amount of enzyme required to release one µmole of inorganic phosphate (P_i) per min under the assay conditions.

An improved phytase plate assay was developed which eliminated false positive results caused by microbial acid production. Bacterial isolates were grown

under anaerobic conditions on modified Scott and Dehority (1965) agar medium containing 5% (v/v) rumen fluid, 1.8% (w/v) agar and 2.0% (w/v) sodium phytate for 5 d at 37°C. Colonies were washed from the agar surface and the petri plates were flooded with a 2% (w/v) aqueous cobalt chloride solution. After a 5-min incubation at room temperature the cobalt chloride solution was replaced with a freshly prepared solution containing equal volumes of a 6.25% (w/v) aqueous ammonium molybdate solution and 0.42% (w/v) ammonium vanadate solution. Following a 5-min incubation, the ammonium molybdate solution/ammonium vanadate solution was removed and the plates examined for zones of clearing. The effectiveness of this counterstaining technique is demonstrated in Figure 1. Prior to staining, zones of clearing were evident around colonies of phytase-producing *S. ruminantium* JY35 (ATCC 55785) and lactic acid-producing *S. bovis* grown on agar medium containing phytate (Figure 1, left petri plate). The false positive zones of clearing resulting from acid production by *S. bovis* colonies were eliminated by counterstaining the plates with cobalt chloride and ammonium molybdate/ammonium vanadate solutions (Figure 1, right petri plate).

B. Phytase activity of ruminal bacteria

The phytase activities of 345 rumen bacteria from the LRC culture collection were determined (Table 1). The anaerobic technique of Hungate (1950), as modified by Bryant and Burkey (1953), or an anaerobic chamber with a 90% CO_2 and 10% H_2 atmosphere was used to cultivate the microorganisms in the LRC culture collection. Phytase screening was performed on isolates grown anaerobically (100% CO_2) in Hungate tubes with 5 mL of modified Scott and Dehority medium (1965) containing 5% (v/v) rumen fluid, 0.2% (w/v) glucose, 0.2% (w/v) cellobiose and 0.3% (w/v) starch. After 18 to 24 h incubation at 39°C, whole cells or culture supernatants were assayed for phytase activity. Selenomonads were the predominant phytase producers (93% of the isolates tested had phytase activity, Table 1). *Prevotella* was the only other genus from which a significant number of positive cultures was identified (11 phytase positive isolates out of 40 tested). A total of 29 cultures with substantial phytase activity were identified. These included 24 of the genus *Selenomonas* and 5 of the genus *Prevotella*. Twelve of these cultures (11

Selenomonas and 1 *Prevotella* isolate) had phytase activities substantially higher than the other positive cultures (Table 2). In all instances, the phytase activity was predominantly cell associated.

Example 3

Phytase activity of *Selenomonas ruminantium* JY35 (ATCC 55785)

A. Growth and phytase production

Phytase production during growth of *S. ruminantium* JY35 (ATCC 55785) was examined. *S. ruminantium* JY35 (ATCC 55785) was grown at 39°C in Hungate tubes with 5 mL of modified Scott and Dehority broth (1965) containing 5% (v/v) ruminal fluid. Growth (protein concentration) and phytase activity (cell associated) were monitored at intervals over a 24-h time period. Maximal growth and phytase activity of *S. ruminantium* JY35 (ATCC 55785) were achieved 8-10 h after inoculation (Figure 2). Cell growth was mirrored by increases in phytase activity.

B. Localization of phytase activity

S. ruminantium JY35 (ATCC 55785) phytase activity was determined to be predominantly cell associated. Little phytase activity was detected in culture supernatants and cell washes. The phytase activity of *S. ruminantium* JY35 (ATCC 55785) was localized by electron microscopy as described by Cheng and Costerton (1973). Cells were harvested by centrifugation, washed with buffer, embedded in 4% (w/v) agar, prefixed in 0.5% glutaraldehyde solution for 30 min and fixed for 2 hours in 5% (v/v) glutaraldehyde solution. Samples were washed five times with cacodylate buffer (0.1 M, pH 7.2) and treated with 2% (w/v) osmium tetroxide, washed five times with cacodylate buffer, dehydrated in a graded ethanol series, and embedded in Spurr's resin (J. B. EM Services Inc.). Ultrathin sections were cut with a Reichert model OM U3 ultramicrotome and stained with 2% (w/v) uranyl acetate and lead citrate. Specimens were viewed with Hitachi H-500 TEM at an accelerating voltage of 75 kV. A comparison of *S. ruminantium* JY35 (ATCC 55785) cells incubated with substrate for reaction product deposition with untreated cells clearly indicated that the phytase activity was associated with the cell outer membrane

surfaces (Figure 3). Deposition of electron dense material on the outer cell surfaces of treated cells was the result of phytase activity (Figures 3A, B and C).

C. Phytase pH optimum

Initial determinations of the pH optimum of the *S. ruminantium* JY35 (ATCC 55785) phytase were conducted with whole cells. Phytase activity was optimal over a pH range of 4.0 to 5.5 (Figure 4). A second pH curve was generated with a $MgCl_2$ cell extract (Figure 5). Cells from a 100-mL overnight culture were washed twice with sterile distilled water, resuspended in 0.3 volumes of a 0.2 M $MgCl_2$ aqueous solution and incubated overnight at 0°C. The solution was clarified by centrifugation and the resulting extract was used in phytase assays. Four buffers systems were used to cover the pH range; glycine (pH 1.5 - 3.0), formate (pH 3.0 - 4.0), acetate (pH 4.0 - 5.5) and succinate (pH 5.5 - 6.5).

D. Phytase temperature optimum

The temperature optimum of the *S. ruminantium* JY35 (ATCC 55785) phytase activity was determined at pH 5.0 (0.1 M sodium acetate buffer) with $MgCl_2$ cell extract. The enzyme retained over 50% of its activity over a temperature range of 37 to 55°C (Figure 6).

E. The effect of ions and substrate concentration on phytase activity

The effect of various ions (10 mM) and substrate concentration on whole cell phytase activity were determined at pH 5.0 (0.1 M sodium acetate buffer). Phytase activity was stimulated by the addition of Ca^{++}, Na^{+}, K^{+} and Mg^{++}, inhibited by Fe^{++}, Zn^{++} and Mn^{++} and unaffected by Co^{++} and Ni^{++} (Figure 7). The effect of substrate concentration on phytase activity in a *S. ruminantium* JY35 (ATCC 55785) $MgCl_2$ cell extract is presented in Figure 8.

F. Molecular Weight

The molecular size of the phytase in *S. ruminantium* JY35 (ATCC 55785) was determined by zymogram analysis. A ten-fold concentrated crude $MgCl_2$ released extract was mixed with 20 µL of sample loading buffer (Laemmli, 1970) in a

microtube and the microtube was placed in a boiling water bath for 5 minutes. The denatured $MgCl_2$ extracts were resolved by SDS-PAGE on a 10% separating gel topped with a 4% stacking gel (Laemmli, 1970). Following electrophoresis, the phytase was renatured by soaking the gel in 1% Triton X-100 for 1 h at room temperature and 0.1 M sodium acetate buffer (pH 5.0) for 1 h at 4°C. Phytase activity was detected by incubating the gel for 16 h in a 0.1 M sodium acetate buffer (pH 5.0) containing 0.4% sodium phytate. The gel was treated with the cobalt chloride and ammonium molybdate/ammonium vanadate staining procedure described for the phytase plate assays in Example 2. A single dominant activity band, corresponding to a molecular mass of approximately 35 to 45 kDa, was observed (Figure 9).

Example 4

Cloning of a phytase gene (*phyA*) from *Selenomonas ruminantium* JY35 (ATCC 55785)

A. Isolation of phytase positive *Escherichia coli* clone

Genomic DNA libraries were prepared for *S. ruminantium* JY35 (ATCC 55785) according to published procedures (Hu *et al.*, 1991; Sambrook *et al.*, 1989). Genomic DNA was extracted from a fresh overnight culture of *S. ruminantium* JY35 (ATCC 55785) using a modification of the protocol described by Priefer *et al.* (1984). *S. ruminantium* JY35 (ATCC 55785) genomic DNA was partially digested with *Sau*3A and gel purified to produce DNA fragments in the 2- to 10-kb range. A genomic library was constructed by ligating *Bam*HI-digested, dephosphorylated pUC18 with *S. ruminantium* JY35 (ATCC 55785) *Sau*3A genomic DNA fragments. *Escherichia coli* DH5α competent cells (Gibco BRL, Mississauga, ON) were transformed with the ligation mix and 6,000 clones carrying inserts were screened for phytase activity (zones of clearing) on LB phytase screening agar [LB medium, 1.0 % sodium phytate (filter sterilized), 100 mM HEPES (pH 6.0 - 6.5), and 0.2 % $CaCl_2$] containing ampicillin (100 µg/mL). A phytase-positive clone SrP.2 was isolated and phytase activity confirmed through enzyme assays (Figure 10). Very high levels of phytase activity were found in the medium as well as associated with the *E. coli* cells (Table

3). Plasmid DNA isolated from clone SrP.2 carried a 5.3-kb plasmid, designated pSrP.2, consisting of pUC18 containing a 2.7-kb *Sau*3A insert.

B. Confirmation of the *Selenomonas ruminantium* JY35 (ATCC 55785) origin of the 2.7-kb insert

The *S. ruminantium* JY35 (ATCC 55785) origin of the 2.7-kb insert in pSrP.2 was confirmed by Southern blot hybridization (Sambrook *et al.*, 1989). Genomic DNA isolated from *S. ruminantium* JY35 (ATCC 55785) and digested with *Eco*RI or *Hin*dIII was resolved on a 0.8% agarose gel. After transfer to Zeta-probe® membrane (BioRad Laboratories), the hybridization was performed overnight at high stringency (2 x SSC; 65°C) with the 2.7-kb fragment from pSrP.2 labelled with digoxigenin (DIG DNA labeling and detection kit; Boehringer Mannheim Canada Ltd., Laval, PQ). The blots were washed twice in 2 x SSC at room temperature; 0.1% SDS for 5 minutes and twice 0.1 x SSC; 0.1% SDS for 20 minutes at 65°C. The blots were developed according to the protocol provided with the DIG DNA labeling and detection kit (Boehringer Mannheim Canada Ltd).

The probe reacted with a 14-kb *Hin*dIII (Figure 11) and a 23-kb *Eco*RI (data not shown) fragment of genomic DNA and confirmed that the 2.7-kb fragment was from *S. ruminantium* JY35 (ATCC 55785) and that a single homologous sequence exists in the genome. Single copies of a sequence homologous to the 2.7-kb fragment from *S. ruminantium* JY35 (ATCC 55785) also exist in the genomes of *S. ruminantium* HD86, HD141, and HD_4 (data not shown). However restriction fragment length polymorphisms were noted for *S. ruminantium* HD86 (9- and 23-kb *Eco*RI fragments) and *S. ruminantium* HD_4 (3-kb *Eco*RI fragment and a 20-kb *Hin*dIII fragment). The labelled 2.7-kb fragment from pSrP.2 failed to hybridize with genomic DNA isolated from *Prevotella* sp. 46/5[2], *E. coli* DH5α or *A. ficuum* NRRL 3135 (data not shown).

Example 5

Characterization of *Selenomonas ruminantium* phytase gene

A. Evidence for the cloning of a phytase gene

Escherichia coli DH5α competent cells (Gibco BRL, Mississauga, ON) were transformed with plasmids pUC18 and pSrP.2. The resulting ampicillin-resistant transformants were tested for phytase activity on LB phytase screening agar. Only *E. coli* DH5α cells transformed with pSrP.2 produced clearing zones on LB phytase screening agar.

B. Restriction and deletion analysis of pSrP.2

The phytase gene was localized on the 2.7-kb *Sau*3A insert by restriction endonuclease and deletion analyses (Ausubel *et al.*, 1990; Sambrook *et al.*, 1989). Cells carrying plasmid pSrP.2Δ*Sph*I, constructed by the deletion of the 1.4-kb *Sph*I fragment from pSrP.2, lacked phytase activity (Figure 12 and Figure 13, Table 3).

C. Zymogram analysis

The molecular mass of the phytase produced by *E. coli* DH5α (pSrP.2) was determined by zymogram analysis. One mL of an overnight culture was transferred to a 1.5-mL microtube. The cells were harvested by centrifugation and washed with 0.1 M sodium acetate buffer (pH 5.5). The cell pellet was resuspended in 80 µL of sample loading buffer (Laemmli, 1970) and the microtube was placed in a boiling water bath for 5 minutes. The resulting cell extracts were resolved by SDS-PAGE on a 10% separating gel topped with a 4% stacking gel (Laemmli, 1970) and the gel was stained for phytase activity as described in Example 3F. A single dominant activity band, corresponding to a molecular mass of approximately 37 kDa, was observed (Figure 14, lane A). A corresponding activity band was not observed for *E. coli* DH5α (pSrP.2Δ*Sph*I) cells (Figure 14, lane B).

D. DNA sequence analysis of pSrP.2

The complete sequence of the 2.7-kb insert of pSrP.2 was determined. Samples were prepared for DNA sequence analysis on an Applied Biosystems Model 373A DNA sequencing system (Applied Biosystems, Inc., Mississauga, ON)

by using a Taq DyeDeoxy™ Terminator Cycle Sequencing Kit (Applied Biosystems, Inc.). Template DNA was extracted from overnight cultures of *E. coli* DH5α (pSrp.2) with the Wizards™ minipreps DNA purification system (Promega Corp., Madison, WI). Overlapping sequences were generated by primer walking. The DNA sequence data was analyzed using MacDNASIS DNA software (Hitachi Software Engineering Co., Ltd., San Bruno, CA).

The sequence of the 2.7-kb DNA insert was determined and DNA structural analysis identified an open reading frame (ORF2; bp 1493 to 2504) overlapping the *Sph*I site of the 2.7-kb *Sau*3A insert and large enough to encode the 37 kDa phytase. Phytase activity was eliminated by deleting bp 1518 through to the end of the 2.7-kb *Sau*3A fragment (pSrPr6, Table 3, Figure 13). This was accomplished by cloning the PCR product of pSrP.2 bounded by sequencing primer SrPr6 (CGG GAT GCT TCT GCC AGT AT, SEQ ID NO. 3 the reverse complement of bp 1518 to 1538) and M13 Forward primer (CGC CAG GGT TTT CCC AGT CAC GAC) into pGEM-T (Promega Corp.). A PCR product subclone (pSrPf6) of pSrP.2, bounded by primer SrPf6 (bp 1232 to 1252, CGT CCA CGG AGT CAC CCT AC) SEQ ID NO. 4 and M13 Reverse primer (AGC GGA TAA CAA TTT CAC ACA GGA), and containing ORF2 plus 252 bp upstream of the *Sph*I cleavage site retained phytase activity (Table 3, Figure 13).

The sequence and translation of the *S. ruminatium* phytase gene (*phyA*) is shown in Figure 15. Translation of ORF2 would result in the expression of a 346-amino acid polypeptide with a predicted molecular weight of 39.6 kDa (Figure 15). The first 31 residues were typical of a prokaryote signal sequence, encompassing a basic N-terminus and central hydrophobic core (von Heijne, 1986). Application of the method of von Heijne (1986) predicted the signal peptidase cleavage site most probably occurs before Ala^{28} or Pro^{31}. This was confirmed by determining the N-terminal amino acid sequence of gel purified from *E. coli* DH5α (pSrPf6) culture supernatant (Figure 15). The secreted mature protein has a putative mass of 36.5 kDa.

A comparison of the *phyA* amino acid sequence with known protein sequences from the MasDNASIS SWISSPROT database revealed no significant similarities to any published sequences including *Aspergillus niger* phytase genes *phyA* and *phyB*.

Example 6

Partial purification and characterization of *phyA* products expressed by *E. coli.*

Cell free supernatants, prepared from overnight cultures of *E. coli* (pSrPf6), were mixed 3:1 (v/v) with Ni^{++}-NTA agarose pre-equilibrated in 0.1 M Tris (pH 7.9), 0.3 M NaCl buffer. The mixture was incubated at room temperature for 0.5 h and washed 3 x with 0.1 M Tris (pH 7.9), 0.3 M NaCl buffer. The phytase activity was eluted from the resin with 1 volume 0.1 M sodium acetate (pH 5.0), 0.3 M NaCl. When resolved on SDS-polyacrylamide gels stained with Coomassie brilliant blue, over 70% of the eluted protein formed a single 37-kDa protein band. Zymogram and N-terminal amino acid sequence analyses confirmed that the 37-kDa band corresponded to the phytase encoded by the cloned *S. ruminantium* JY35 (ATCC 55785) *phy*A. The specific activity of Ni^{++}-NTA agarose-purified phytase ranged from 200 to 400 µmol phosphate released/min/mg protein. This is 2 to 4 times higher than the specific activity reported for the purified *A. ficuum* NRRL 3135 phytase (van Gorcum et al., 1991, 1995; van Hartingsveldt et al., 1993).

Example 7

Overexpression of the *Selenomonas ruminatium phy*A gene

Isolation and characterization of *phy*A from *S. ruminantium* JY35 (ATCC 55785) enables the large scale production of protein PhyA in any of a number of prokaryotic (e.g., *E. coli* and *B. subtilis*) or eukaryotic (e.g., fungal - *Pichia, Saccharomyces, Aspergillus, Trichoderma*; plant - *Brassica, Zea, Solanum*; or animal - poultry, swine or fish) expression systems using known methods. Teachings for the construction and expression of *phy*A in *E. coli*, *P. pastoris*, and *B. napus* are provided below. Similar approaches may be adopted for expression of the *S. ruminantium* JY35 (ATCC 55785) phytase in other prokaryotic and eukaryotic organisms.

A. Cloning of the *Selenomonas ruminatium phyA* in an *Escherichia coli* - specific expression construct

An expression construct is constructed in which the region encoding the mature PhyA is transcriptionally fused with the *tac* promoter (Brosius et al., 1985). The promoter sequences may be replaced by those from other promoters that provide for efficient expression in *E. coli*. The expression construct is introduced into *E. coli* cells by transformation.

i. Construction of the *E. coli* expression vector

A number of *E. coli* expression vectors based on the *tac* or related promoters are commercially available. In this example the construct will be prepared with pKK223-3 available from Pharmacia Biotech Inc. (Uppsala, Sweden). The region of *phyA* encoding the mature PhyA (the peptide secreted following removal of the signal peptide) is amplified with oligonucleotide primers MATE2 (GC GAA TTC ATG GCC AAG GCG CCG GAG CAG AC) (SEQ ID NO. 5) and M13 Reverse. The oligonucleotide MATE2 (SEQ ID NO. 5) was designed to contain a suitable restriction site at its terminus to allow direct assembly of the amplified product with pKK223-3. The region of *phyA* amplified with MATE2 (SEQ ID NO. 5) and M13 Reverse is digested with *Eco*RI and *Sma*I and ligated into similarly cleaved pKK223-3.

ii. Transformation of *E. coli* and PhyA expression

The pKK223-3::*phyA* ligation mix is used to transform competent *E. coli* cells. Strains suitable for high levels of protein expression, such as SG13009, CAG926 or CAG929 (carrying *lacI* on a plasmid such as pREP4), are employed. Transformed cells are spread on LB agar containing ampicillin (100 µg/mL) and incubated overnight at 37°C. Ampicillin-resistant colonies are screened for the presence of the desired pKK223-3::*phyA* construct by extracting pDNA and subjecting the pDNA to agarose gel electrophoresis and restriction analysis. Positive clones may be further characterized by PCR and DNA sequence analysis.

Expression of the *S. ruminantium* JY35 (ATCC 55785) phytase by transformed *E. coli* cells is tested by growing the cells under vigorous aeration at 37°C in a suitable liquid medium (e.g., LB or 2xYT) containing the appropriate antibiotic selection until the optical density (at 600 nm) is between 0.5 and 1.0. The *tac* promoter is induced by adding isopropyl-β-D-thiogalactoside (IPTG) to a final

concentration between 0.1 and 2 mM. The cells are cultivated for an additional 2 to 4 h and harvested by centrifugation. Protein expression is monitored by SDS-PAGE, and western blot/immunodetection techniques. The expressed PhyA may be extracted by breaking (e.g., sonication or mechanical disruption) the *E. coli* cells. Protein inclusions of PhyA may be harvested by centrifugation and solubilized with 1 to 2 % SDS. The SDS may be removed by dialysis, electroelution or ultrafiltration. The phytase activity of prepared cell extracts may be assayed by standard methods described in Example 2.

B. Cloning of the *Selenomonas ruminatium* phyA in a *Pichia pastoris* - specific expression construct

An expression construct is constructed in which the region encoding the mature PhyA is translationally fused with the secretion signal sequences found on *P. pastoris* expression vectors (Pichia Expression Kit Instruction Manual, Invitrogen Corporation, San Diego, CA) in order to express the *S. ruminantium* phytase as a secreted product. The promoter and secretion signal sequences may be replaced by those from other promoters that provide for efficient expression in *Pichia*. The expression construct is introduced into *P. pastoris* cells by transformation.

i. Construction of the *P. pastoris* expression vector

A number of *P. pastoris* expression vectors based on the *aox1* promoters and α- Factor or *pho*1 signal sequences are commercially available. In this example the construct will be prepared with pPIC9 available from Invitrogen Corporation. The region of *phyA* encoding the mature PhyA is amplified with oligonucleotide primers MATE (GC GAA TTC GCC AAG GCG CCG GAG CAG AC) (SEQ ID NO. 6) and M13 Reverse. The oligo MATE (SEQ ID NO. 6) was designed to contain a suitable restriction site at its terminus to allow direct assembly of the amplified product with pPIC9. The region of *phyA* amplified with MATE (SEQ ID NO. 6) and M13 Reverse is digested with *Eco*RI and ligated into similarly cleaved pPIC9.

ii. Transformation of *P. pastoris* and PhyA expression

The pPIC9::*phyA* ligation mix is used to transform competent *E. coli* DH5α cells. Transformed cells are spread on LB agar containing ampicillin (100 μg/mL) and incubated overnight at 37°C. Ampicillin-resistant colonies are screened for the

presence of the desired pPIC9::*phy*A construct by extracting pDNA and subjecting the pDNA to agarose gel electrophoresis and restriction analysis. Positive clones are further characterized by PCR and DNA sequence analysis. Plasmid DNA is prepared from a 1 L culture of an *E. coli* clone carrying the desired pPIC9::*phy*A construct. The pDNA is digested with *Bgl*II and analyzed by agarose gel electrophoresis to confirm complete digestion of the vector. The digested pDNA is extracted with phenol:chloroform, ethanol precipitated and resuspended in sterile distilled H_2O to a final concentration of 1 µg/mL. In preparation for transformation, *P. pastoris* GS115 or KM71 cells are grown for 24 h at 30°C in YPD broth. Cells from 100 µL of culture are harvested by centrifugation and resuspended in 100 µL of transformation buffer (0.1M LiCl, 0.1M dithiothreitol, 45% polyethylene glycol 4000) containing 10 µg salmon sperm DNA and 10 µg of linearized pPIC9::*phy*A. The mixture is incubated for 1 h at 37°C, spread on *P. pastoris* minimal agar medium and incubated for 2 to 5 d. Colonies growing on the minimal agar medium are streaked for purity and analyzed for the presence of the integrated *phy*A by PCR and Southern blot hybridization.

Expression of the *S. ruminantium* JY35 (ATCC 55785) phytase by transformed *P. pastoris* cells is tested by growing the cells at 30°C under vigorous aeration in a suitable liquid medium (e.g. buffered complex glycerol medium such as BMGY) until a culture optical density (at 600 nm) (OD_{600}) of 2 to 6 is reached. The cells are harvested and resuspended to an OD_{600} of 1.0 in an inducing medium (e.g., buffered complex methanol medium, BMMY) and incubated for a further 3 to 5 days. Cells and cell-free culture supernatant are collected and protein expression is monitored by enzyme assay, SDS-PAGE, and western blot/immunodetection techniques.

C. Cloning of the *Selenomonas ruminatium phyA* in a *Pichia pastoris* - specific expression construct - A Further Example

An expression construct is constructed in which the region encoding the mature PhyA is translationally fused with the secretion signal sequences found on *P. pastoris* expression vectors (e.g., *Pichia* Expression Kit Instruction Manual, Invitrogen Corporation, San Diego, CA) in order to express the *S. ruminantium*

phytase as a secreted product. The promoter and secretion signal sequences may be replaced by those from other promoters that provide for efficient expression in *Pichia*. The expression construct is introduced into *P. pastoris* cells by transformation.

i. Construction of the *P. pastoris* expression vector

A number of *P. pastoris* expression vectors based on the *aox1* promoters and α-Factor or *pho1* signal sequences are commercially available. In this example the construct was prepared with pPICZαA available from Invitrogen Corporation. The region of *phyA* encoding the mature PhyA (i.e., the peptide secreted following removal of the signal peptide) was amplified with oligonucleotide primers MATE (GC GAA TTC GCC AAG GCG CCG GAG CAG AC SEQ ID NO. 6) and M13 Reverse. The oligo MATE (SEQ ID NO. 6) was designed to contain an *Eco*RI restriction site at its terminus to allow direct assembly of the amplified product with pPICZαA. The region of *phyA* amplified with MATE (SEQ ID NO. 6) and M13 Reverse was digested with *Eco*RI and ligated into similarly cleaved pPICZαA.

ii. Transformation of *P. pastoris*

The pPICZαA::*phyA* ligation mix was used to transform competent *E. coli* DH5α cells. Transformed cells were spread on LB agar containing Zeocin (25 mg/mL) and incubated overnight at 37°C. Zeocin resistant colonies were screened for the presence of the desired pPICZαA::*phyA* construct by extracting pDNA and subjecting the pDNA to agarose gel electrophoresis and restriction analysis. Positive clones were further characterized by PCR and DNA sequence analysis. Plasmid DNA was prepared from a 1 L culture of an *E. coli* clone carrying the desired pPICZαA::*phyA* construct. The pDNA is digested with *Bgl*II and analyzed by agarose gel electrophoresis to confirm complete digestion of the vector. The digested pDNA was extracted with phenol:chloroform, ethanol precipitated and resuspended in sterile distilled H_2O to a final concentration of 1 μg/μL.

In preparation for transformation, 50 mL of YPD broth were inoculated with *P. pastoris* GS115 cells and incubated at 28°C and 250 RPM for 1 day. Subsequently, 5 mL of the 1 d culture was used to inoculate 50 mL of fresh YPD broth. The culture was propagated overnight at 28°C and 250 RPM. The following morning, 5 mL of this culture was used to inoculate 50 mL of fresh YPD broth. This culture was

incubated at 28°C and 250 RPM until the culture OD_{600} reached approximately 1.2 (~ 6 h). The yeast cells from 20 ml of fresh culture were harvest by centrifugation, washed once with and resuspended in 1 mL of room temperature 10 mM Tris, 1 mM EDTA, 0.1 M LiCl, 0.1 M dithiothreitol buffer (pH 7.4). After a 1 h incubation at 30°C, the cell suspension was washed once with 1 mL ice cold water and once with 1 mL ice cold 1 M sorbitol. The cells were resuspended in 160 µL of ice cold 1 M sorbitol (to obtain cell concentrations approaching 10^{10} cells/mL). Linearized pPICZαA::*phyA* (5 to 10 µg) was mixed with 80 µL of cells, loaded into prechilled electroporation cuvettes (0.2 cm inter-electrode distance) and incubated on ice for 5 min. A high voltage pulse (1.5 kV, 25 µF, 200 Ohms) was applied to the cuvette with a Bio-Rad Gene Pulser™. Immediately following the pulse, 1 mL of ice cold 1M sorbitol was added to the cuvette which was incubated subsequently for 2 h at 30°C. The cell suspension was spread (100 to 200 µL per plate) on YPD agar medium containing Zeocin (100 µg/mL) and incubated for 2 to 4 d at 30°C. Colonies growing on the selective medium were streaked for purity and analyzed for the presence of the integrated *phyA* by PCR and/or Southern blot hybridization.

iii. *Pichia pastoris* expression of the *S. ruminantium* JY35 phytase gene

Expression of the *S. ruminantium* JY35 phytase by transformed *P. pastoris* cells was tested by growing transformed cells grown overnight in buffered complex glycerol medium (e.g., buffered complex glycerol medium, BMGY, *Pichia* Expression Kit Instruction Manual) at 28°C and 250 RPM and transferring them into inducing medium (e.g., buffered complex methanol medium, BMMY). The cells harvested from the BMGY medium were washed once with BMMY medium, resuspended in BMMY to an OD_{600} of 1.0 and incubated for a further 3 to 5 days at 28°C and 250 RPM. Methanol (0.005 volumes) was added every 24 h. Cells and cell free culture supernatants were collected and assayed for phytase activity.

Sixteen *P. pastoris* pPICZαA::MATE transformants were tested for phytase activity following 96 h growth in BMMY medium. The most active transformant, named clone 17, was selected for further study. Growth and phytase production by *P. pastoris* pPICZαA::MATE clone 17 and a negative clone (*P. pastoris* pPICZαA) were monitored over a period of 9 d. Starter cultures were prepared by growing the isolates overnight (28°C, 250 RPM) in 10 mL of BMGY (glycerol) medium. The cells

were harvested and duplicate cultures were prepared by resuspending the cells in 50 mL BMMY (methanol) medium to an approximate OD_{600} of 2.5. The resulting cultures were transferred into 500 mL flasks and incubated at 28°C and 250 RPM. Methanol was added every 24 h to a final concentration of 0.5%. Optical density and phytase activity were measured over the time course of the experiment. The results are presented in Table 4. Phytase activity was detected only in cultures carrying the *S. ruminantium phyA* gene. These cultures produced up to 22.5 units of phytase activity per mL after 210.5 h cultivation.

Phytase activity in shake flask cultures was increased through modification of the induction protocol and medium composition. The phytase activity of clone 17 was dramatically improved by increasing the initial cell density (OD_{610} = 36.0) of the induced culture. After nearly 4 d growth (91.5 h), phytase activities greater than 40 and 20 units/mL were observed for whole culture and cell free supernatant samples, respectively. The optical densities (OD_{610}) of these cultures were between 62 and 69. Experimental results suggest that the greater the culture biomass at the time of methanol induction, the greater the yields of recombinant phytase. Biomass yields as high as 150 g/L (dry weight) or optical densities of 1500 have been reported for *Pichia* cultivated under optimal growth conditions in a tightly controlled fermentor system operating with oxygen enrichment.

Pichia phytase yields were also increased by adding Tween-80 to the medium. Surfactants have been shown previously to affect phytase production by *Aspergillus carbonarius* (Al-Asheh and Duvnjak, 1994). The effect of incorporating 0, 0.02, 0.1 or 0.5 % Tween-80 on phytase yields of BMMY cultures of *P. pastoris* pPICZαA::MATE clone 17 is illustrated in Table 5. The cells from 2 d YPD cultures were harvested and resuspended in BMMY (OD_{610} = 8.3). Triplicate flasks for each concentration of Tween-80 were prepared and incubated at 28°C and 250 RPM. Methanol (0.005 volumes) was added on a daily basis to the flasks. Phytase activity increased more rapidly in cultures containing higher concentrations of Tween-80. Furthermore, a larger proportion of the phytase activity was found in the supernatant when higher Tween-80 concentrations were used. Phytase yields as high as 298 units/mL of shake flask culture have been achieved with a 9 d culture of clone 17 cultivated in BMMY medium amended with 0.5% Tween-80.

Cellular and supernatant proteins were analyzed by SDS-PAGE to confirm the production of PhyA by *P. pastoris*. The presence of a 37 kDa protein band was readily apparent when as little as 5 µL of supernatant was resolved on a 12% SDS-PAGE gel. The 37 kDa band was visible in the cellular protein sample but represented less than 10% of that found in the corresponding amount of supernatant. In addition to PhyA, supernatants from clone 17 contained very few additional proteins (a useful characteristic of *Pichia* expression). The recombinant PhyA protein comprised over 95% (estimated from SDS-PAGE gels) of the secreted protein. The 37 kDa protein band was not present in the supernatant or cells of a negative control culture (*P. pastoris* pPICZαA).

Shake flask experiments with recombinant *P. pastoris* cells expressing the *S. ruminantium* phytase (PhyA) have demonstrated the potential of this protein production system. Significant gains in phytase yields will be obtained by cultivating and inducing clone 17 in a fermentor. Additional gains in phytase yields may be achieved by increasing gene copy number through further screening of independent transformants or the use of multicopy vector systems. Spontaneous multiple plasmid integration events occur in *Pichia* at a frequency between 1/10 and 1/100 transformants. It is not unrealistic to expect that a 10 fold gain in phytase yield (e.g., 3,000 units/mL) may be readily achieved through manipulation of phytase gene copy number and control of fermentation parameters. This would result in production levels comparable to commercial *A. ficuum* phytase production systems. Yields for these systems are believed to be around 3,000,000 units (µmol Pi released/min) of phytase activity per L of culture.

iv. The Activity of recombinant the *S. ruminantium* phytase (PhyA) on grain substrates

The liberation of phosphate from corn by the recombinant *S. ruminantium* JY35 phytase produced by *Pichia pastoris* was examined. Feed corn was ground and sieved through a mesh to obtain a particle size between 1 - 3 mm. Ground corn (0.5 g) was weighed into sterile 15 mL Falcon tubes to which 2 mL of 0.1 M sodium acetate buffer (pH 5.0) was added. After addition of phytase, the reaction mixtures were incubated at 37°C. Phosphate release was determined by measuring supernatant phosphate. In order to measure the background phosphate, reaction

mixtures were prepared and terminated immediately through the addition of 5% (w/v) TCA. All experiments were conducted in triplicate.

Incubation of corn in a sodium acetate buffer resulted in the release of increasing amounts of phosphorus over time (Table 6). Although the addition of phytase activity significantly increased the amount of phosphorus released, the rate of phosphorus release decreased with time.

The concentration of phytase added to the incubation mixture also influenced the amount of phosphorus released. Raising phytase concentrations from 0.08 units to 0.48 units per g of corn resulted in increased levels of phosphorus in the supernatant (Table 7). It should be noted that increasing the phytase concentration from 0.32 to 0.48 units produced only a marginal increase in phosphorus released.

D. Cloning of the *Selenomonas ruminatium* phyA in a *Brassica napus* seed - specific expression construct

Transformation and gene expression methods have been developed for a wide variety of monocotyledonous and dicotyledonous crop species. In this example, a *S. ruminantium* JY35 (ATCC 55785) phytase expression construct is constructed in which the region encoding the mature PhyA is translationally fused with an oleosin coding sequence in order to target seed oil body specific expression of the *S. ruminantium* phytase. The promoter and/or secretion signal sequences may be replaced by those from other promoters that provide for efficient expression in *B. napus* or other transformable plant species. The expression construct is introduced into *B. napus* cells by *Agrobacterium*-mediated transformation.

i. Construction of the *B. napus* expression vector

A number of expression vectors functional in *B. napus* are described in the literature (Gelvin et al., 1993). In this example, the construct is prepared by replacing the *E. coli* β-glucuronidase CDS of pCGOBPGUS (van Rooijen and Moloney, 1994) with a fragment encoding the *phyA* mature CDS. This is accomplished by subcloning the pCGOBPGUS *PstI KpnI* fragment, containing the oleosin promoter::oleosin CDS::β-glucuronidase CDS::NOS region, on to *PstI KpnI*- digested pUCBM20 (Boehringer Mannheim Canada, Laval, PQ). This plasmid is called pBMOBPGUS. The region of *phyA* encoding the mature PhyA is amplified with

oligonucleotide primers MATN (GA GGA TCC ATG GCC AAG GCG CCG GAG CAG AC) (SEQ ID NO. 7) and M13 Reverse. The oligonucleotide MATN (SEQ ID NO. 7) was designed to contain a suitable restriction site at its terminus to allow direct assembly of the amplified product with digested pBMOBPGUS. The *phy*A fragment amplified with MATN (SEQ ID NO. 7) and M13 Reverse is digested with *Nco*I *Sst*I and ligated into similarly cleaved pBMOBPGUS to generate plasmid pBMOBP*phy*A. The *B. napus* expression vector, pCGOBP*phy*A, is constructed by replacing the *Pst*I *Kpn*I fragment from pCGOBPGUS with the *Pst*I *Kpn*I fragment from pBMOBP*phy*A, containing the oleosin promoter::oleosin CDS::phyA CDS::NOS fragment.

ii. Transformation of *B. napus* and stable PhyA expression

Transgenic *B. napus* is prepared as described by van Rooijen and Moloney (1994). *Agrobacterium tumefaciens* strain EHA101 is transformed by electroporation with pCGOBP*phy*A. Cotyledonary petioles of *B. napus* are transformed with *A. tumefaciens* EHA101 (pCGOBP*phy*A). Transgenic plants are regenerated from explants that root on hormone-free MS medium containing 20 µg/mL kanamycin. Young plants are assayed for NPTII activity, grown to maturity and allowed to self pollenate and set seed. Seeds from individual transformants are pooled and part of the seed sample is assayed for the presence of phytase activity and compared to seeds from untransformed plants. Second generation plants (T2) are propagated from the seeds of clones with the highest levels of phytase activity. Seeds from the T2 plants homozygous for NPTII (hence also for *phy*A) are selected and used for mass propagation of plants (T3) capable of producing the highest amounts of phytase.

Example 8

Identification of Related Phytase Genes in Other Microorganisms

To identify a phytase gene related to phyA, hybridization analysis can be used to screen nucleic acids from one or more ruminal isolates of interest using *phy*A (SEQ ID NO. 1) or portions thereof as probes by known techniques (Sambrook, 1989; Ausubel, 1990) as described in example 4B. Related nucleic acids may be cloned by employing known techniques. Radioisotopes (i.e., ^{32}P) may be required

when screening organisms with complex genomes in order to incr ase the sensitivity of the analysis. Polymerase chain reaction (PCR) amplification may also be used to identify genes related to *phyA*. Related sequences found in pure or mixed cultures are preferentially amplified by PCR (and variations of such as Reverse Transcription - PCR) with oligonucleotides primers designed using SEQ ID NO. 1. Amplified products may be visualized by agarose gel electrophoresis and cloned using known techniques. A variety of materials, including cells, colonies, plaques, and extracted nucleic acids (e.g., DNA, RNA), may be examined by these techniques for the presence of related sequences. Alternatively, known immunodetection techniques employing antibodies specific to PhyA (SEQ ID NO. 2) can be used to screen whole cells or extracted proteins of interest for the presence of related phytase(s).

Table 1. Phytase activity among rumen bacteria.

Phytase Activity	Microorganism	Number of isolates tested
Very Strong	*Prevotella* sp.	1
	Selenomonas ruminantium	11
Strong	*Prevotella ruminicola*	4
	S. ruminantium	13
Moderate	*Bacillus* sp.	1
	Megasphaera elsdenii	7
	P. ruminicola	6
	S. ruminantium	37
	Treponema sp.	1
Negative	*Anaerovibrio lipolytica*	2
	Bacillus sp.	4
	Butyrivibrio fibrisolvens	47
	Clostridium sp.	1
	Coprococcus sp.	3
	Enterococcus sp.	4
	Eubacterium sp.	7
	Fibrobacter succinogenes	8
	Fusobacterium sp.	3
	Lachnospira multiparus	4
	Lactobacillus sp.	20
	M. elsdenii	7
	Peptostreptococcus sp.	1
	P. ruminicola	41
	Ruminobacter amylophilus	4
	Ruminococcus albus	7
	Ruminococcus flavefaciens	10
	S. ruminantium	4
	Streptococcus bovis	48
	Streptococcus milleri	1
	Staphylococcus sp.	6
	Succinovibrio dextrisolvens	12
	Treponema sp.	12
	Unknown	8
	Total isolates screened	345

Table 2. Phytase activity of selected rumen bacterial isolates.

Isolate	Phytase activity (mU*/mL)
Selenomonas ruminantium JY35	646
Selenomonas ruminantium KJ118	485
Selenomonas ruminantium BS131	460
Selenomonas ruminantium HD141	361
Selenomonas ruminantium HD86	286
Selenomonas ruminantium JY135	215
Selenomonas ruminantium D	69
Selenomonas ruminantium HD16	52
Selenomonas ruminantium BS114	47
Selenomonas ruminantium JY4	27
Prevotella sp. 46/5[2]	321
Prevotella ruminicola JY97	68
Prevotella ruminicola KJ182	61
Prevotella ruminicola JY106	49
Megasphaera elsdenii JY91	5

*U= μmoles, P_i released/min

Table 3. Overexpression of *S. ruminantium*[1] phytase in recombinant *E.coli* DH5α.

Strain	Sample Composition	Units[2] /mL	Specific Activity (Units/mg protein)
E. coli (pSrP.2)	cells	0.30 (0.08)[3]	1.56 (0.41)
	supernatant	0.308 (0.21)	2.64 (1.51)
E. coli (pSrPf6)	cells	0.91 (0.41)	6.42 (0.64)
	supernatant	5.10 (0.58)	22.83 (1.67)
E. coli (pSrP.2 *Sph*I)	cells	ND[4]	ND
	supernatant	ND	ND

[1] *S. ruminantium* JY35 is a crescent shaped-rod, an obligate anaerobe, produces proprionic acid from the fermentation of glucose, ferments lactose, does not ferment glycerol, does not ferment mannitol (see also Bergey's Manual of Systematic Bacteriology, ed. John G. Holt, Williams and Wilkins, Baltimore, 1984)

[2] Units = μmoles P_i released/min

[3] Numbers in parenthese are standard errors

[4] ND = not detected

Table 4. Growth and phytase activity of *P. pastoris* cells transformed with pPICZαA (negative control) or pPICZαA::MATE (clone 17).

Culture	Time (h)	Optical Density (610 nm)	Phytase activity (μmol/min/mL) Culture	Supernatant
P. pastoris (pPICZαA)	0.0	2.6	0.0	0.0
	20.5	10.1	0.0	0.0
	42.5	17.8	0.0	0.0
	68.0	17.0	0.0	0.0
	91.0	28.5	0.0	0.0
	138.5	39.3	0.0	0.0
	210.5	46.7	0.0	0.0
P. pastoris (pPICZαA::MATE)	0.0	2.5	0.0	0.0
	20.5	11.3	1.9	0.1
	42.5	13.9	4.4	1.5
	68.0	12.9	8.0	2.7
	91.0	15.7	4.7	0.5
	138.5	18.3	12.6	5.3
	210.5	18.7	22.5	12.5

Table 5. The effect of Tween-80 concentration on growth and phytase activity of *P. pastoris* cells transformed with pPICZαA::MATE (clone 17).

Time (d)	Sample (% Tween-80)	Optical Density (610 nm)	Phytase Activity (μmol/min/mL) Culture	Phytase Activity (μmol/min/mL) Supernatant	Supernatant/ Culture Activity
2	0.0	24.3	4.1	2.2	0.55
	0.02	24.4	4.8	2.7	0.57
	0.1	25.1	5.2	3.2	0.61
	0.5	24.4	4.9	3.2	0.65
4	0.0	31.2	6.9	4.7	0.69
	0.02	31.0	8.2	5.5	0.67
	0.1	31.8	10.3	6.9	0.67
	0.5	29.2	10.3	9.1	0.88
8	0.0	32.8	10.6	5.9	0.55
	0.02	30.4	14.8	9.8	0.67
	0.1	33.9	20.2	17.2	0.86
	0.5	33.8	22.1	18.9	0.86

Table 6. The effect of incubation period and recombinant *S. ruminantium* JY35 phytase (2 units/g of corn) on phosphate release from corn.

Sample	Length of incubation (h)	Phosphate concentration (µmoles/mL)
No phytase	1	0.85
	2	1.72
	3	2.56
	4	3.77
	5	4.35
Phytase	1	4.76
	2	6.83
	3	7.72
	4	8.41
	5	8.49

Table 7. The effect of recombinant *S. ruminantium* JY35 phytase concentration on phosphate release from corn.

Phytase activity (units/g of corn)	Phosphate concentration (µmoles/g of corn)
0.08	11.8
0.16	14.8
0.24	22.5
0.32	23.0
0.40	23.2
0.48	23.8
0.56	23.8
0.64	23.6
0.72	23.8

REFERENCES

Al-Asheh, S. And Z. Duvnjak. 1994. The effect of surfactants on the phytase production and the reduction of the phytic content in canola meal by *Aspergillus carbonarius* during a solid state fermentation process. Biotechnol. Lett. 16:183-188.

Ausubel, F.A., R. Brent, R.E. Kingston, D.D. Moore, J.G. Sneidman, J.A. Smith and K. Struhl. (eds.) 1990. Current protocols in molecular biology. Green Publishing and Wiley-Interscience, New York.

Brosius, J., M. Erfle and J. Storella. 1985. Spacing of the -10 and -35 regions in the *tac* promoter. J. Biol. Chem. 260:3539-3541.

Bryant, M.P. and L.A. Burkey. 1953. Cultural methods and some characteristics of some of the numerous groups of bacteria in the bovine rumen. J. Dairy Sci. 36:205-217.

Cheng, E.W., G. Hall and W. Burroughs. 1955. A method for the study of cellulose digestion by washed suspensions of rumen microorganisms. J. Dairy Sci. 38:1255-1230.

Cheng, K.-J. and J.W. Costerton, 1973. Localization of alkaline phosphatase in three Gram-negative rumen bacteria. J. Bacteriol. 116:424-440.

Dayhoff, M.O., R.M. Schwartz and B.C. Orcutt. 1978. A model of evoluntionary change in proteins. In: Atlas of Protein Sequence and Structure. Volume 5, Supplement 3, Chapter 22, pp 345-352.

Ellis, S.B., P.F. Brust, P.J. Koutz, A.F. Waters, M.M. Harpold and R.R. Gingeras. 1985. Isolation of alcohol oxidase and two other methanol regulated genes from the yeast, *Pichia pastoris*. Mol. Cell. Biol. 5:1111-1121.

Fiske, C.H. and Y. Subbarow. 1925. The colorimetric determination of phosphorus. J. Biol. Chem. 66:376-400.

Gelvin, S.B., R.A. Schilperoort and D.P.S. Verma (eds.). 1993. Plant Molecular Biology Manual. Kluwer Academic Publishers, Boston, MA.

Graf, E. (ed.). 1986. Phytic acid, chemistry and applications. Pilatus Press. Minneapolis, MN. 344 pp.

Howson, S.J. and R.P. Davis. 1983. Production of phytate-hydrolysing enzyme by some fungi. Enzyme Microb. Technol. 5:377-382.

Hu, Y.J., D.C. Smith, K.-J. Cheng and C.W. Forsberg. 1991. Cloning of a xylanase gene from *Fibrobacter succinogenes* 135 and its expression in *Escherichia coli*. Can.J. Microbiol. 37:554-561.

Hungate, R.E. 1950. The anaerobic mesophilic cellulolytic bacteria. Bacteriol. Rev. 14:1-49.

Laemmli, U.K. 1970. Cleavage of the structural proteins during assembly of the head of bacteriophage T4. Nature 227:680-685.

Priefer, U., R. Simon and A. Puhler. 1984. Cloning with cosmids. *In*: Puhler, A. and K.N. Timmis (eds) Advanced molecular genetics. Springer-Verlag, New York. pp.190-201.

Raun, A., E. Cheng and W. Burroughs. 1956. Phytate phosphorus hydrolysis and availability to rumen microorganisms. Agric. Food Chem. 4:869-871.

Sambrook, J., E.F. Fritsch and T. Maniatis. 1989. Molecular cloning. A laboratory manual. 2nd. edn. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY.

Scott, H.W. and B.A. Dehority. 1965. Vitamin requirements of several cellulolytic bacteria. J. Bacteriol. 89:1169-1175.

Shieh, T.R. and J.H. Ware. 1968. Survey of microorganisms for the production of extracellular phytase. Appl. Microbiol. 16:1348-1351.

van Gorcom, R.F.M. and C.A.M.J Van Den Hondel. 1993. Cloning, characterization and overexpression of the phytase gene (*phyA*) of *Aspergillus niger*. Gene 127:87-94.

van Hartingsveldt W., C.M.J. van Zeij, M.G. Harteveld, R.J. Gouka, M.E.G. Suykerbuyk, R.G.M. Luiten, P.A. Van Paridon, G.C.M. Selten, A.E. Veenstra, van Rooijen, G.J.H. and M. M. Moloney. 1994. Plant seed oil-bodies as carriers for foreign proteins. Bio/Technology 13:72-77.

von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. 14:4683-4690.

Wong, S.-L. 1989. Development of an inducible and enhancible expression and secretion system in *Bacillus subtilis*. Gene 83:215-223.

All publications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains. All publications are herein incorporated by reference to the same extent as if each individual publication was specifically indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity and understanding, it will be obvious that certain changes and modifications may be practised within the scope of the claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Cheng, Kuo-Joan
Selinger, Leonard B.
Yanke, Lindsey J.
Bae, Hee-Dong
Zhou, Lu Ming
Forsberg, Cecil W.

(ii) TITLE OF INVENTION: DNA sequences encoding phytases of ruminal microorganisms.

(iii) NUMBER OF SEQUENCES: 7

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: McKay-Carey & Company
(B) STREET: 2125, 10155-102 St.
(C) CITY: Edmonton
(D) STATE: Alberta
(E) COUNTRY: CA
(F) ZIP: T5J 4G8

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: May 23, 1997
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Mary Jane McKay-Carey
(B) REGISTRATION NUMBER:
(C) REFERENCE/DOCKET NUMBER: 37003WO0

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (403) 424-0222
(B) TELEFAX: (403) 421-0834

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1401 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Selenomonas ruminantium
(B) STRAIN: JY35

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Genomic DNA library
(B) CLONE: pSrP.2

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 231..1268
(C) IDENTIFICATION METHOD: experimental
(D) OTHER INFORMATION: /codon_start= 231
/function= "Dephosphorylation of phytic acid"
/product= "Phytase"
/evidence= EXPERIMENTAL
/gene= "phyA"
/number= 1
/standard_name= "myo-inositol hexaphosphate phosphohydrolase"
/citation= ([1])

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 231..311
(C) IDENTIFICATION METHOD: experimental
(D) OTHER INFORMATION: /codon_start= 1
/function= "phytase secretion"
/product= "Signal peptide"
/evidence= EXPERIMENTAL
/citation= ([1])

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 312..1268
(C) IDENTIFICATION METHOD: experimental
(D) OTHER INFORMATION: /codon_start= 312
/product= "Phytase"
/evidence= EXPERIMENTAL
/number= 2
/citation= ([1])

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGTCCACGGA GTCACCCTAC TATACGACGT ATGTGAAGTT CACGTCGAAG TTCTAGGGAA 60

TCACCGATTC GTGCAGGATT TTACCACTTC CTGTTGAAGC GGATGAGAAG GGGAACCGCG 120

AAGCGGTGGA AGAGGTGCTG CACGACGGAC GATCGCGCTG AATGAATCAG TGCTTCCTAA 180

```
CTATTGGGAT TCCGCGCAGA CGCGCGGATG GAGTAAAGGA GTAAGTTGTT ATG AAA          236
                                                        Met Lys
                                                        -27

TAC TGG CAG AAG CAT GCC GTT CTT TGT AGT CTC TTG GTC GGC GCA TCC          284
Tyr Trp Gln Lys His Ala Val Leu Cys Ser Leu Leu Val Gly Ala Ser
-25                 -20                 -15                 -10

CTC TGG ATA CTG CCG CAG GCC GAT GCG GCC AAG GCG CCG GAG CAG ACG          332
Leu Trp Ile Leu Pro Gln Ala Asp Ala Ala Lys Ala Pro Glu Gln Thr
                -5                  1                   5

GTG ACG GAG CCC GTT GGG AGC TAC GCG CGC GCG GAG CGG CCG CAG GAC          380
Val Thr Glu Pro Val Gly Ser Tyr Ala Arg Ala Glu Arg Pro Gln Asp
        10                  15                  20

TTC GAG GGC TTT GTC TGG CGC CTC GAC AAC GAC GGC AAG GAG GCG TTG          428
Phe Glu Gly Phe Val Trp Arg Leu Asp Asn Asp Gly Lys Glu Ala Leu
    25                  30                  35

CCG CGT AAT TTC CGC ACG TCG GCT GAC GCG CTG CGC GCG CCG GAG AAG          476
Pro Arg Asn Phe Arg Thr Ser Ala Asp Ala Leu Arg Ala Pro Glu Lys
40                  45                  50                  55

AAA TTC CAT CTC GAC GCC GCG TAT GTA CCG TCG CGC GAG GGC ATG GAT          524
Lys Phe His Leu Asp Ala Ala Tyr Val Pro Ser Arg Glu Gly Met Asp
                60                  65                  70

GCA CTC CAT ATC TCG GGC AGT TCC GCA TTC ACG CCG GCG CAG CTC AAG          572
Ala Leu His Ile Ser Gly Ser Ser Ala Phe Thr Pro Ala Gln Leu Lys
            75                  80                  85

AAC GTT GCC GCG AAG CTG CGG GAG AAG ACG GCT GGC CCC ATC TAC GAT          620
Asn Val Ala Ala Lys Leu Arg Glu Lys Thr Ala Gly Pro Ile Tyr Asp
        90                  95                  100

GTC GAC CTA CGG CAG GAG TCG CAC GGC TAT CTC GAC GGT ATC CCC GTG          668
Val Asp Leu Arg Gln Glu Ser His Gly Tyr Leu Asp Gly Ile Pro Val
    105                 110                 115

AGC TGG TAC GGC GAG CGC GAC TGG GCA AAT CTC GGC AAG AGC CAG CAT          716
Ser Trp Tyr Gly Glu Arg Asp Trp Ala Asn Leu Gly Lys Ser Gln His
120                 125                 130                 135

GAG GCG CTC GCC GAC GAG CGG CAC CGC TTG CAC GCA GCG CTC CAT AAG          764
Glu Ala Leu Ala Asp Glu Arg His Arg Leu His Ala Ala Leu His Lys
                140                 145                 150

ACG GTC TAC ATC GCG CCG CTC GGC AAG CAC AAG CTC CCC GAG GGC GGC          812
Thr Val Tyr Ile Ala Pro Leu Gly Lys His Lys Leu Pro Glu Gly Gly
            155                 160                 165

GAA GTC CGC CGC GTA CAG AAG GTG CAG ACG GAA CAG GAA GTC GCC GAG          860
Glu Val Arg Arg Val Gln Lys Val Gln Thr Glu Gln Glu Val Ala Glu
        170                 175                 180
```

```
GCC GCG GGG ATG CGC TAT TTC CGC ATC GCG GCG ACG GAT CAT GTC TGG        908
Ala Ala Gly Met Arg Tyr Phe Arg Ile Ala Ala Thr Asp His Val Trp
    185                 190                 195

CCA ACG CCG GAG AAC ATC GAC CGC TTC CTC GCG TTT TAC CGC ACG CTG        956
Pro Thr Pro Glu Asn Ile Asp Arg Phe Leu Ala Phe Tyr Arg Thr Leu
200                 205                 210                 215

CCG CAG GAT GCG TGG CTC CAT TTC CAT TGT GAA GCC GGT GTC GGC CGC       1004
Pro Gln Asp Ala Trp Leu His Phe His Cys Glu Ala Gly Val Gly Arg
                220                 225                 230

ACG ACG GCG TTC ATG GTC ATG ACG GAT ATG CTG AAG AAC CCG TCC GTA       1052
Thr Thr Ala Phe Met Val Met Thr Asp Met Leu Lys Asn Pro Ser Val
            235                 240                 245

TCG CTC AAG GAC ATC CTC TAT CGC CAG CAC GAG ATC GGC GGC TTT TAC       1100
Ser Leu Lys Asp Ile Leu Tyr Arg Gln His Glu Ile Gly Gly Phe Tyr
        250                 255                 260

TAC GGG GAG TTC CCC ATC AAG ACG AAG GAT AAA GAT AGC TGG AAG ACG       1148
Tyr Gly Glu Phe Pro Ile Lys Thr Lys Asp Lys Asp Ser Trp Lys Thr
    265                 270                 275

AAA TAT TAT AGG GAA AAG ATC GTG ATG ATC GAG CAG TTC TAC CGC TAT       1196
Lys Tyr Tyr Arg Glu Lys Ile Val Met Ile Glu Gln Phe Tyr Arg Tyr
280                 285                 290                 295

GTG CAG GAG AAC CGC GCG GAT GGC TAC CAG ACG CCG TGG TCG GTC TGG       1244
Val Gln Glu Asn Arg Ala Asp Gly Tyr Gln Thr Pro Trp Ser Val Trp
                300                 305                 310

CTC AAG AGC CAT CCG GCG AAG GCG TAAAAGCGCA GGCGGCGGCT CGGAGTCAGG      1298
Leu Lys Ser His Pro Ala Lys Ala
            315

GAAATGGCGC TGCCAGCACG GGACGCGCGG CGGCGGATGC TGCGCCGGTC AGGGATGATT     1358

GACGACAGCC AGAGAAGAAA GGATGGTTTT ATGAGGTGGA TCC                       1401
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 346 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Lys Tyr Trp Gln Lys His Ala Val Leu Cys Ser Leu Leu Val Gly
-27      -25                 -20                 -15
```

```
Ala Ser Leu Trp Ile Leu Pro Gln Ala Asp Ala Ala Lys Ala Pro Glu
-10                 -5                  1                   5

Gln Thr Val Thr Glu Pro Val Gly Ser Tyr Ala Arg Ala Glu Arg Pro
                10                  15                  20

Gln Asp Phe Glu Gly Phe Val Trp Arg Leu Asp Asn Asp Gly Lys Glu
            25                  30                  35

Ala Leu Pro Arg Asn Phe Arg Thr Ser Ala Asp Ala Leu Arg Ala Pro
        40                  45                  50

Glu Lys Lys Phe His Leu Asp Ala Ala Tyr Val Pro Ser Arg Glu Gly
    55                  60                  65

Met Asp Ala Leu His Ile Ser Gly Ser Ser Ala Phe Thr Pro Ala Gln
70                  75                  80                  85

Leu Lys Asn Val Ala Ala Lys Leu Arg Glu Lys Thr Ala Gly Pro Ile
                90                  95                  100

Tyr Asp Val Asp Leu Arg Gln Glu Ser His Gly Tyr Leu Asp Gly Ile
            105                 110                 115

Pro Val Ser Trp Tyr Gly Glu Arg Asp Trp Ala Asn Leu Gly Lys Ser
        120                 125                 130

Gln His Glu Ala Leu Ala Asp Glu Arg His Arg Leu His Ala Ala Leu
    135                 140                 145

His Lys Thr Val Tyr Ile Ala Pro Leu Gly Lys His Lys Leu Pro Glu
150                 155                 160                 165

Gly Gly Glu Val Arg Arg Val Gln Lys Val Gln Thr Glu Gln Glu Val
                170                 175                 180

Ala Glu Ala Ala Gly Met Arg Tyr Phe Arg Ile Ala Ala Thr Asp His
            185                 190                 195

Val Trp Pro Thr Pro Glu Asn Ile Asp Arg Phe Leu Ala Phe Tyr Arg
        200                 205                 210

Thr Leu Pro Gln Asp Ala Trp Leu His Phe His Cys Glu Ala Gly Val
    215                 220                 225

Gly Arg Thr Thr Ala Phe Met Val Met Thr Asp Met Leu Lys Asn Pro
230                 235                 240                 245

Ser Val Ser Leu Lys Asp Ile Leu Tyr Arg Gln His Glu Ile Gly Gly
                250                 255                 260

Phe Tyr Tyr Gly Glu Phe Pro Ile Lys Thr Lys Asp Lys Asp Ser Trp
            265                 270                 275
```

Lys Thr Lys Tyr Tyr Arg Glu Lys Ile Val Met Ile Glu Gln Phe Tyr
280 285 290

Arg Tyr Val Gln Glu Asn Arg Ala Asp Gly Tyr Gln Thr Pro Trp Ser
295 300 305

Val Trp Leu Lys Ser His Pro Ala Lys Ala
310 315

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide SrPr6"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Selenomonas ruminantium
(B) STRAIN: JY35

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Genomic DNA library
(B) CLONE: pSrP.2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CGGGATGCTT CTGCCAGTAT 20

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide SrPf6"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Selenomonas ruminantium

(B) STRAIN: JY35

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Genomic DNA library
(B) CLONE: pSrP.2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGTCCACGGA GTCACCCTAC 20

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide MATE2"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Selenomonas ruminantium
(B) STRAIN: JY35

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Genomic DNA library
(B) CLONE: pSrP.2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCGAATTCAT GGCCAAGGCG CCGGAGCAGA C 31

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide MATE"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Selenomonas ruminantium
(B) STRAIN: JY35

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Genomic DNA library
(B) CLONE: pSrP.2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GCGAATTCGC CAAGGCGCCG GAGCAGAC 28

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide MATN"

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Selenomonas ruminantium
(B) STRAIN: JY35

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Genomic DNA library
(B) CLONE: pSrP.2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAGGATCCAT GGCCAAGGCG CCGGAGCAGA C 31

WE CLAIM:

1. A purified and isolated DNA encoding a phytase of a ruminal microorganism.

2. A purified and isolated DNA according to claim 1 wherein said ruminal microorganism is a prokaryote.

3. A purified and isolated DNA according to claim 1 wherein said ruminal microorganism is of the genus *Selenomonas, Prevotella, Treponema* or *Megasphaera*.

4. A purified and isolated DNA according to claim 1 wherein said ruminal microorganism is *Selenomonas ruminantium, Prevotella ruminicola, Treponema bryantii* or *Megasphaera elsdenii*.

5. A purified and isolated DNA according to claim 1 wherein said ruminal microorganism is *Selenomonas ruminantium*.

6. A purified and isolated DNA according to claim 1 wherein said ruminal microorganism is *Selenomonas ruminantium* JY35 (ATCC 55785).

7. A purified and isolated DNA according to claim 1, said DNA being capable of hybridizing under stringent conditions with a probe comprising at least 25 continuous nucleotides of nucleotide sequence SEQ ID NO. 1.

8. A purified and isolated DNA according to claim 1, said phytase comprising amino acid sequence SEQ ID NO. 2.

9. A purified and isolated DNA according to claim 1, said DNA comprising nucleotide sequence SEQ ID NO. 1.

10. A purified and isolated DNA according to claim 1, said DNA comprising nucleotides 312-1268 of SEQ ID NO. 1.

11. A purified and isolated DNA according to claim 1, wherein said encoded phytase has the following characteristics:

a) a molecular mass of about 37 kDa;

b) is active within a pH range of about 3.0 to 6.0; and

c) is active within a temperature range of about 4 to 55°C.

12. A purified and isolated DNA according to claim 11 wherein said encoded phytase is active within a temperature range of about 20 to 55°C.

13. A purified and isolated DNA according to claim 11 wherein said encoded phytase is active within a temperature range of about 35 to 40°C.

14. A purified and isolated DNA according to claim 11, wherein the encoded phytase has the following additional characteristic:

d) a specific activity at least two fold higher than that of *Aspergillus ficuum* NRRI 3135 PhyA as measured by the release of inorganic phosphate.

15. An expression construct capable of directing the expression of a phytase in a suitable host cell, said expression construct comprising a DNA encoding a phytase of a ruminal microorganism operably linked to control sequences compatible with said host cell.

16. An expression construct according to claim 15 wherein said ruminal microorganism is *Selenomonas ruminantium*.

17. An expression construct according to claim 15 wherein said encoded phytase comprises amino acid sequence SEQ ID NO. 2.

18. A host cell transformed with a DNA encoding a phytase of a ruminal microorganism so that the host cell can express the phytase encoded by said DNA.

19. A transformed host cell according to claim 18 wherein said ruminal microorganism is *Selenomonas ruminantium*.

20. A transformed host cell according to claim 18 wherein said encoded phytase comprises amino acid sequence SEQ ID NO. 2.

21. A transformed host cell according to claim 18 wherein said host cell is a eukaryote.

22. A transformed host cell according to claim 18 wherein said host cell is a prokaryote.

23. A transformed host cell according to claim 18 wherein said host cell is a *Pichea pastoris* cell.

24. A transformed host cell according to claim 18 wherein said host cell is a *Bacillus subtilis* cell.

25. A transformed host cell according to claim 18 wherein said host cell is an *E. coli* cell

26. *Selenomonas ruminantium* JY35 (ATCC 55785).

27. A transgenic plant transformed with a DNA encoding a phytase of a ruminal microorganism so that the phytase encoded by said DNA can be expressed by said plant.

28. A transgenic plant according to claim 27 wherein said ruminal microorganism is *Selenomonas ruminantium*.

29. A transgenic plant according to claim 27 wherein said encoded phytase comprises amino acid sequence SEQ ID NO. 2.

30. A phytase of a ruminal microorganism.

31. A phytase according to claim 30 wherein said ruminal microorganism is *Selenomonas ruminantium*.

32. A phytase according to claim 30 wherein said phytase has the following characteristics:

a) a molecular mass of about 37 kDa;

b) is active within a pH range of about 3.0 to 6.0; and

c) is active within a temperature range of about 4 to 55°C.

33. A phytase according to claim 32 having the following additional characteristic:

d) a specific activity at least two fold higher than that of *Aspergillus ficuum* NRRI 3135 PhyA as measured by the release of inorganic phosphate.

34. A phytase according to claim 30, comprising a contiguous amino acid sequence residing within amino acid sequence SEQ ID NO. 2.

35. A phytase according to claim 30 comprising amino acid sequence SEQ ID NO. 2.

36. A feed composition comprising a feedstuff treated with a phytase of a ruminal microorganism.

37. A feed composition according to claim 36 wherein said ruminal microorganism is *Selenomonas ruminantium*.

38. A feed composition according to claim 36 wherein said phytase comprises amino acid sequence SEQ ID NO. 2.

39. A feed composition according to claim 36 containing a sufficient amount of said phytase to provide up to about 2000 Units (μmoles phosphate released/minute) of phytase activity per kg feed composition.

40. A feed composition according to claim 36 containing a sufficient amount of said phytase to provide up to about 1000 Units of phytase activity per kg feed composition.

41. A feed composition according to claim 36 containing a sufficient amount of said phytase to provide from about 50 to 800 Units of phytase activity per kg feed composition.

42. A feed composition according to claim 36 containing a sufficient amount of said phytase to provide from about 300 to 800 Units of phytase activity per kg feed composition.

43. A feed additive comprising a preparation of lyophilized microorganisms, said microorganisms expressing a phytase of a ruminal microorganism under normal growing conditions.

44. A feed additive according to claim 43 wherein said microorganism is *Selenomonas ruminantium*.

45. A feed additive according to claim 43 wherein said microorganism is a recombinant microorganism transformed with a DNA encoding said phytase of said ruminal microorganism.

46. A feed additive according to claim 45 wherein said ruminal microorganism is *Selenomonas ruminantium*.

47. A feed additive according to claim 45 wherein said expressed phytase comprises amino acid sequence SEQ ID NO. 2.

48. A feed additive for treatment of a feedstuff, said feed additive comprising a phytase of a ruminal microorganism.

49. A feed additive according to claim 48 wherein said microorganism is *Selenomonas ruminantium*.

50. A feed additive according to claim 48 wherein said phytase comprises amino acid sequence SEQ ID NO. 2.

51. A method for producing a phytase, comprising:
(a) transforming at least one host cell with a DNA encoding a phytase of a ruminal microorganism so that said host cell can express said phytase; and
(b) growing a culture of said host cells under conditions conducive to the expression of said phytase by said host cells.

52. A method according to claim 51 comprising the further step of:
(c) extracting said phytase from said culture.

53. A method according to claim 51 wherein said ruminal microorganism is *Selenomonas ruminantium*.

54. A method according to claim 51 wherein said phytase comprises amino acid sequence SEQ ID NO. 2.

55. A method for producing a transgenic plant, comprising:
(a) transforming a plant with a DNA encoding a phytase of a ruminal microorganism so that said plant can express said phytase; and
(b) growing said plant under conditions conducive to the expression of said phytase by said plant.

56. A method according to claim 55 wherein said ruminal microorganism is *Selenomonas ruminantium*.

57. A method according to claim 55 wherein said phytase comprises amino acid sequence SEQ ID NO. 2.

58. A method for improving dietary phytate utilization by an animal, comprising feeding said animal a diet which includes an effective amount of a phytase of a ruminal microorganism.

59. A method according to claim 58 wherein said ruminal microorganism is *Selenomonas ruminantium*.

60. A method according to claim 58 wherein said phytase comprises amino acid sequence SEQ ID NO. 2.

61. A method according to claim 58 wherein said diet includes drinking water and said phytase is included in said drinking water.

62. A method according to claim 58 wherein said phytase is provided in a mineral block for consumption by said animal.

63. A method according to claim 58 wherein said phytase is provided in a pill for consumption by said animal.

64. A method according to claim 58 wherein said phytase is provided in a gel formulation for consumption by said animal.

65. A method according to claim 58 wherein said phytase is sprayed in a liquid formulation onto a feedstuff for consumption by said animal.

66. A method according to claim 58 wherein said phytase is provided in a pelletized feedstuff for consumption by said animal.

67. A method according to claim 58 wherein a feedstuff for consumption by said animal is treated with a preparation of lyophilized microorganisms, said microorganisms expressing said phytase under normal growing conditions.

68. A method for assaying phytase activity of a microorganism, comprising the steps of:

(a) providing a growth medium upon which colonies of microorganisms have been grown, said medium containing a source of phytate;

(b) contacting said medium with an aqueous solution of cobalt chloride; and

(c) examining said medium for zones of clearing,

whereby false positive results caused by microbial acid production are eliminated.

69. A method according to claim 68 wherein after step (b), said medium is contacted with an aqueous solution of ammonium molybdate and an aqueous solution of ammonium vanadate;

70. A method according to claim 68 wherein said medium is contacted with said aqueous solution of cobalt chloride for at least about 5 minutes.

71. A method according to claim 68 wherein said medium is contacted with said aqueous solutions of ammonium molybdate and ammonium vanadate for at least about 5 minutes.

72. A method according to claim 68 wherein said medium is contacted with said aqueous solutions of ammonium molybdate and ammonium vanadate simultaneously.

73. A method according to claim 68 wherein the concentration of said aqueous solution of cobalt chloride is about 2% (weight/volume).

74. A method according to claim 69 wherein the concentration of said aqueous solution of ammonium molybdate is about 6% (weight/volume) and the concentration of said aqueous solution of ammonium vanadate is about 0.5% (weight/volume).

75. A method for identifying a nucleic acid molecule from an organism, said nucleic acid molecule encoding a phytase, said method comprising the steps of:

(a) isolating nucleic acid molecules from said organism;

(b) performing nucleic acid hybridization under conditions of moderate to high stringency with said nucleic acid molecules and a labelled hybridization probe having a nucleotide sequence comprising at least 25 continuous nucleotides of SEQ ID NO: 1.

76. A method according to claim 75 wherein said hybridization conditions are of moderate stringency.

77. A method according to claim 75 wherein said hybridization conditions are of high stringency.

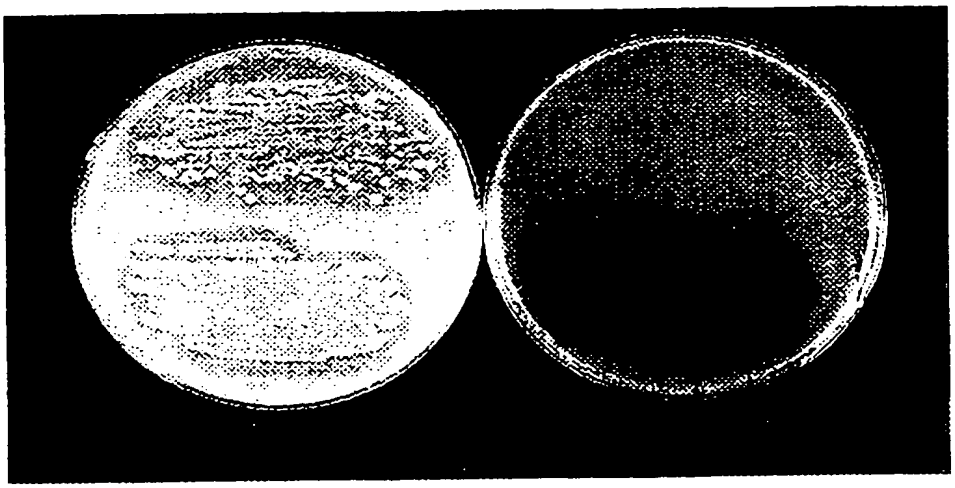


Figure 1

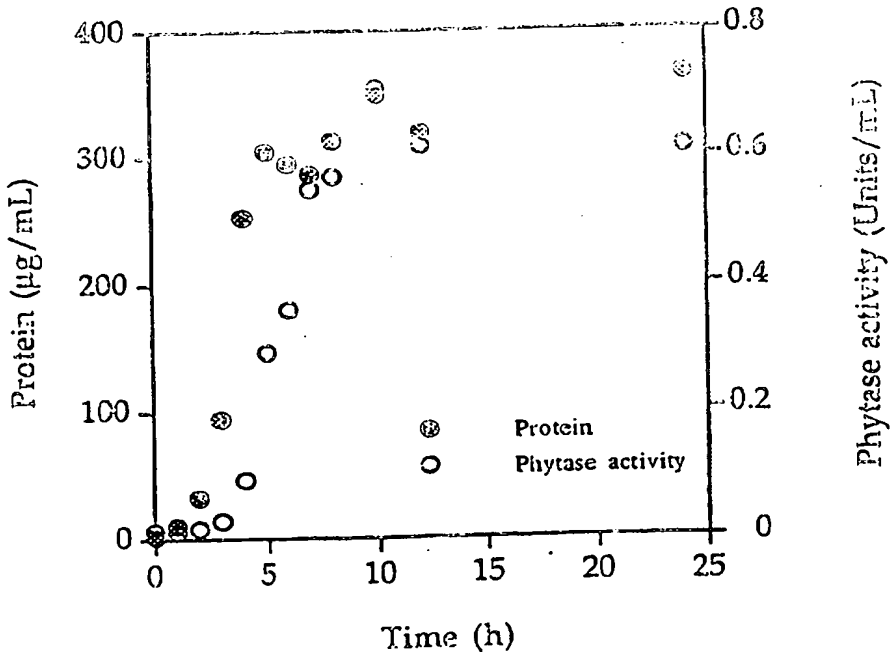


Figure 2

Fig.3A

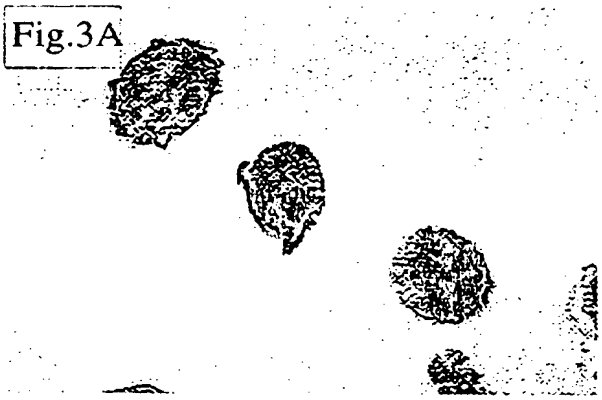


Fig. 3D

Fig.3B

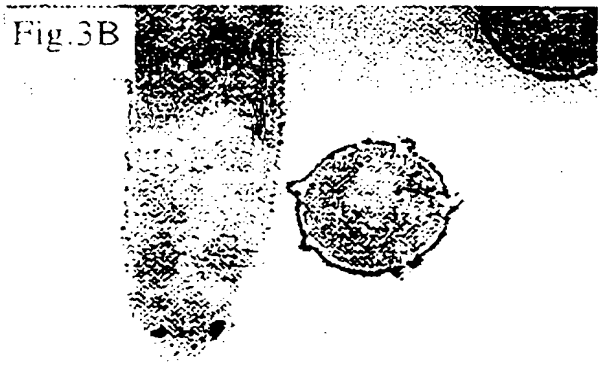


Fig. 3E

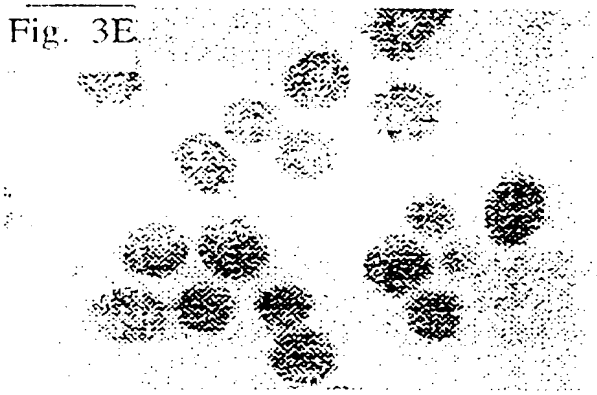


Fig. 3C

Fig. 3F

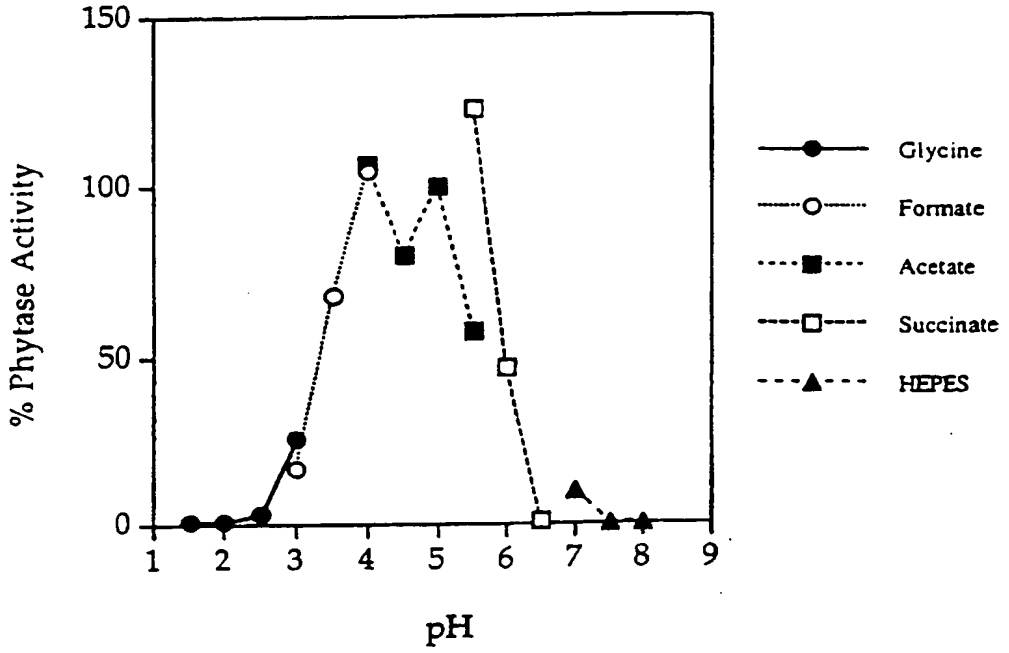


Figure 4

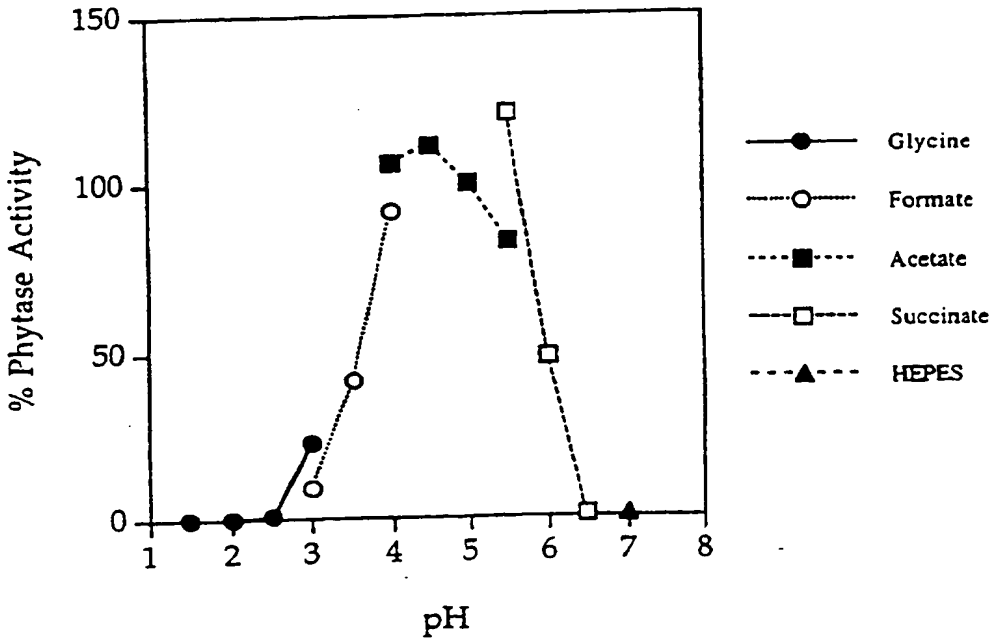


Figure 5

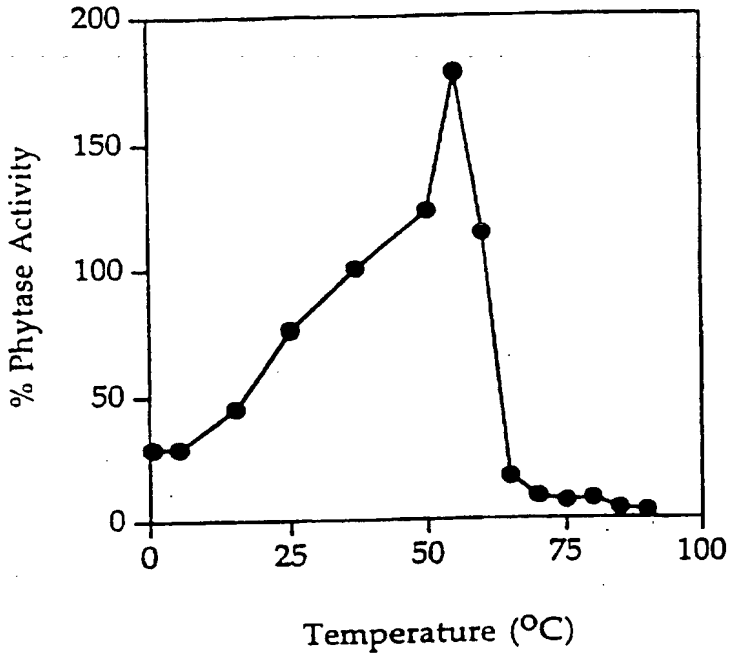


Figure 6

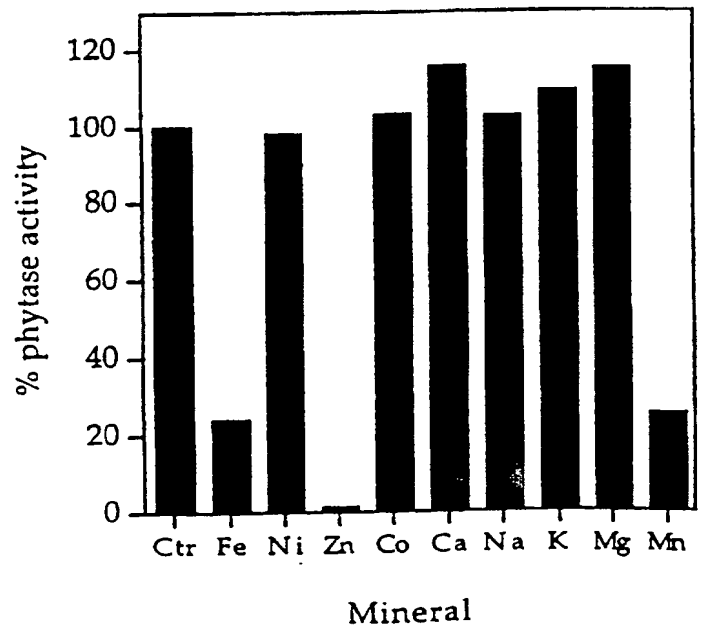


Figure 7

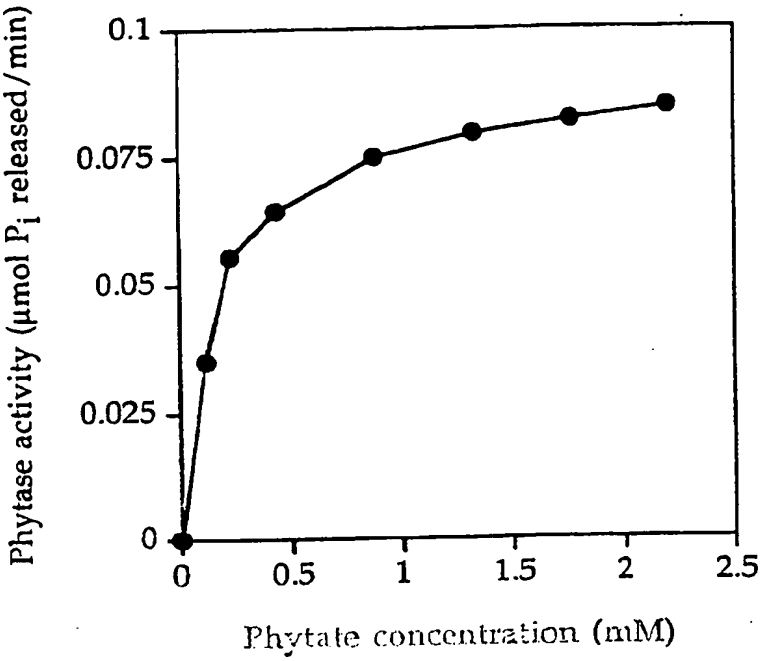


Figure 8

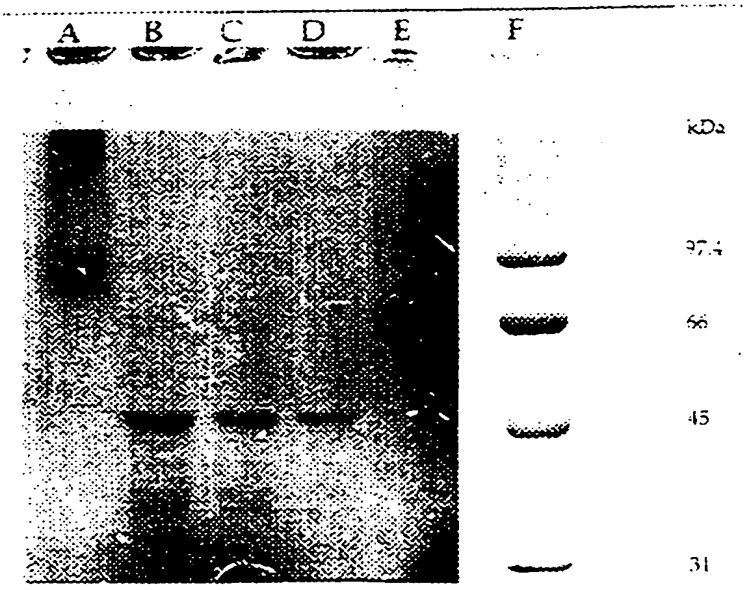


Figure 9

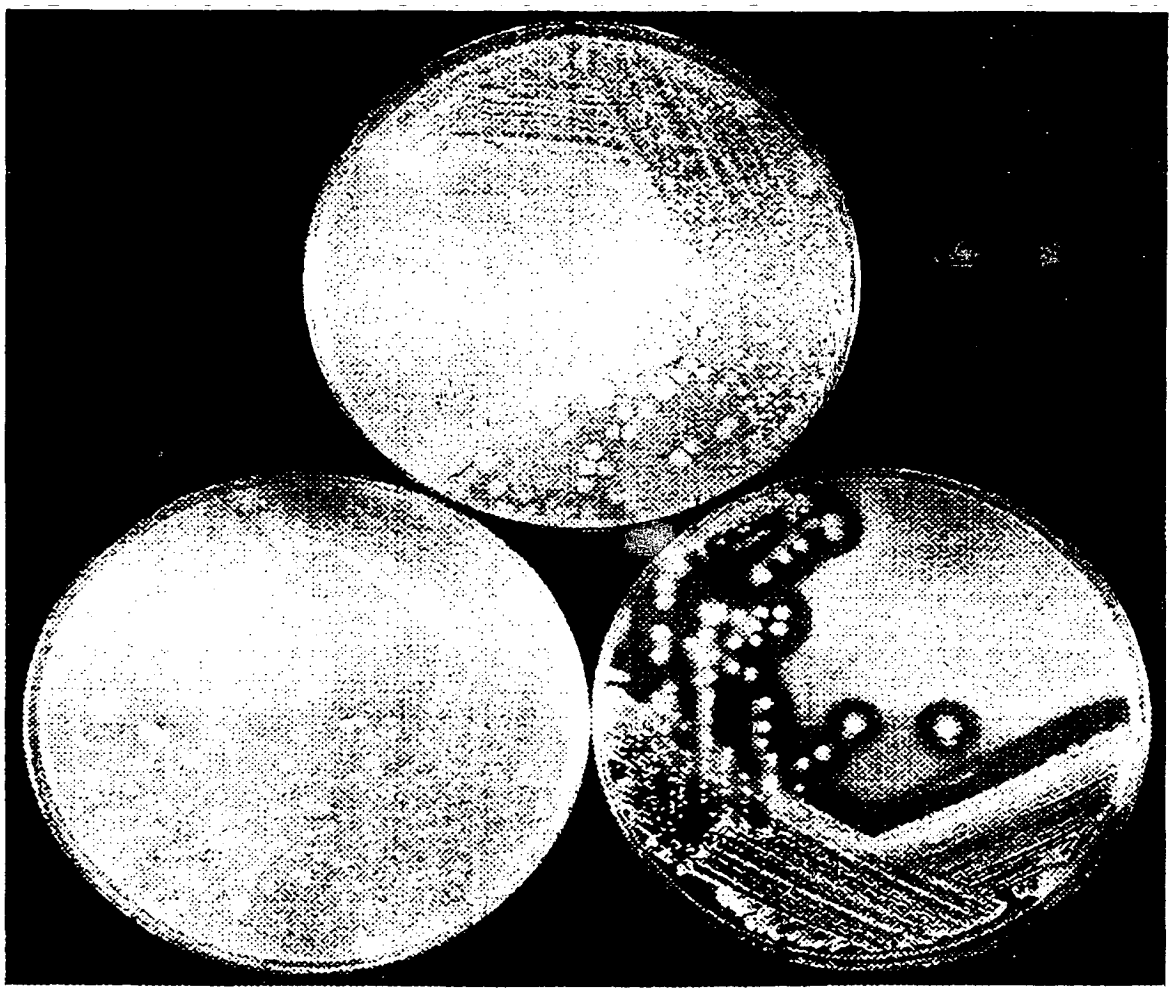


Figure 10

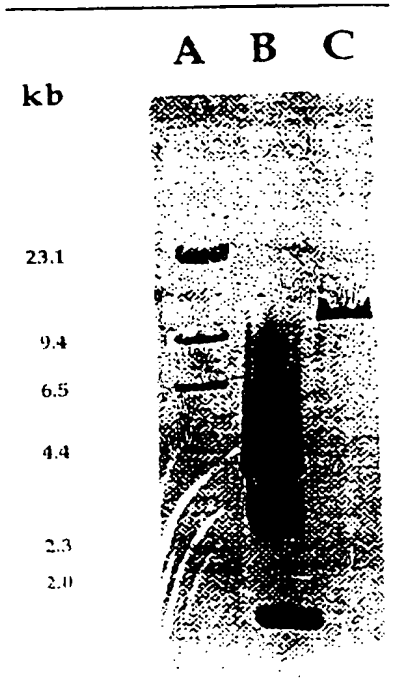


Figure 11

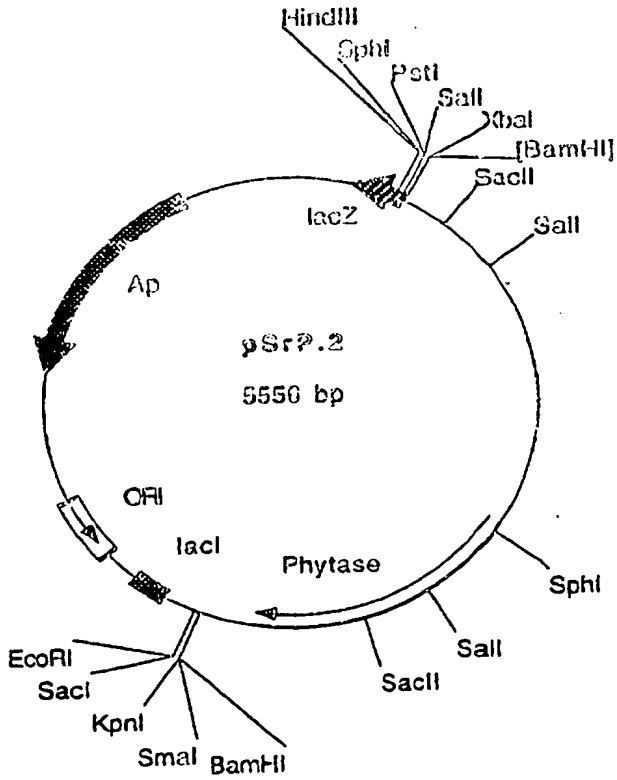


Figure 12

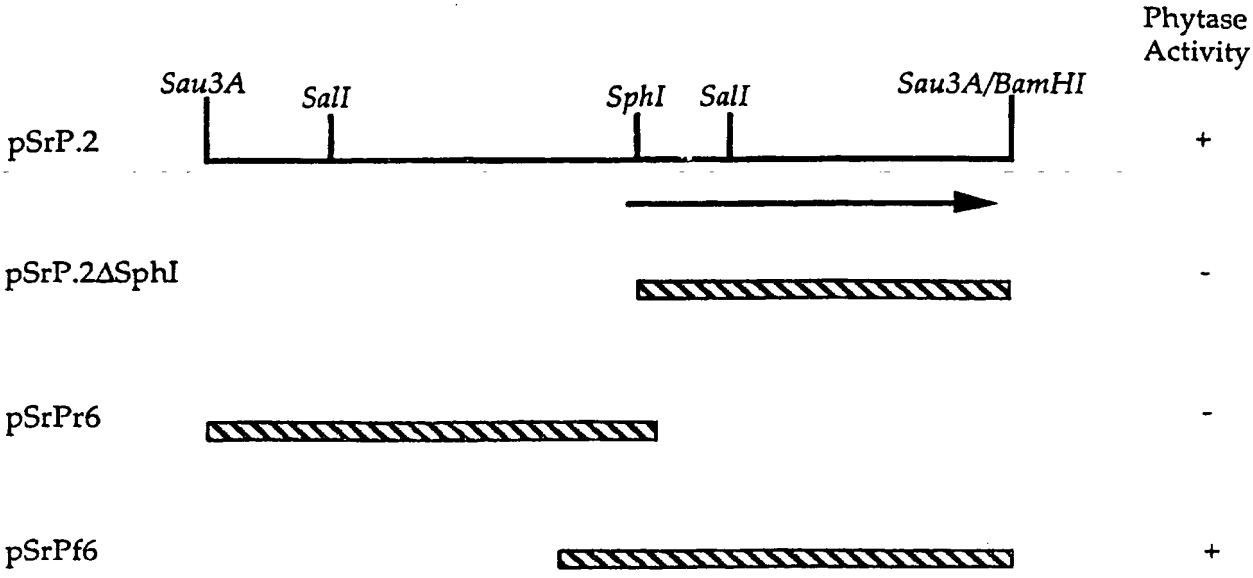


Figure 13

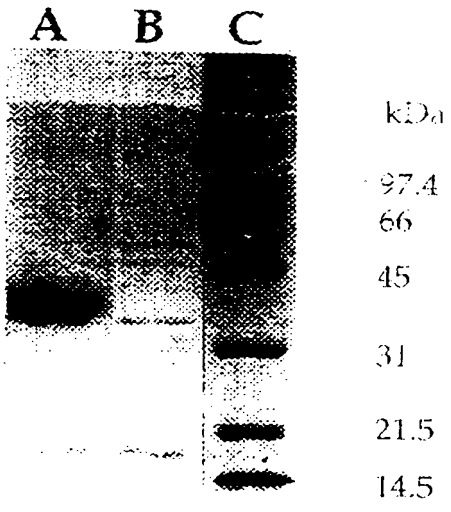


Figure 14

Figure 15

```
1    CGTCCACGGA GTCACCCTAC TATACGACGT ATGTGAAGTT CACGTCGAAG TTCTAGGGAA   60

61   TCACCGATTC GTGCAGGATT TTACCACTTC CTGTTGAAGC GGATGAGAAG GGGAACCGCG  120

121  AAGCGGTGGA AGAGGTGCTG CACGACGGAC GATCGCGCTG AATGAATCAG TGCTTCCTAA  180
                                        R.B.S.
181  CTATTGGGAT TCCGCGCAGA CGCGCGGATG GAGTAAAGGA GTAAGTTGTT ATG AAA TAC  239
                                                              M  K  Y    3
1

240  TGG CAG AAG CAT GCC GTT CTT TGT AGT CTC TTG GTC GGC GCA TCC CTC TGG  290
4    W  Q  K  H  A  V  L  C  S  L  L  V  G  A  S  L  W                 20

291  ATA CTG CCG CAG GCC GAT GCG GCC AAG GCG CCG GAG CAG ACG GTG ACG GAG  341
21   I  L  P  Q  A  D  A  A  K  A  P  E  Q  T  V  T  E                 37
                        ↑
342  CCC GTT GGG AGC TAC GCG CGC GCG GAG CGG CCG CAG GAC TTC GAG GGC TTT  392
38   P  V  G  S  Y  A  R  A  E  R  P  Q  D  F  E  G  F                 54

393  GTC TGG CGC CTC GAC AAC GAC GGC AAG GAG GCG TTG CCG CGT AAT TTC CGC  443
55   V  W  R  L  D  N  D  G  K  E  A  L  P  R  N  F  R                 71

444  ACG TCG GCT GAC GCG CTG CGC GCG CCG GAG AAG AAA TTC CAT CTC GAC GCC  494
72   T  S  A  D  A  L  R  A  P  E  K  K  F  H  L  D  A                 88

495  GCG TAT GTA CCG TCG CGC GAG GGC ATG GAT GCA CTC CAT ATC TCG GGC AGT  545
89   A  Y  V  P  S  R  E  G  M  D  A  L  H  I  S  G  S                105

546  TCC GCA TTC ACG CCG GCG CAG CTC AAG AAC GTT GCC GCG AAG CTG CGG GAG  596
106  S  A  F  T  P  A  Q  L  K  N  V  A  A  K  L  R  E                122

597  AAG ACG GCT GGC CCC ATC TAC GAT GTC GAC CTA CGG CAG GAG TCG CAC GGC  647
123  K  T  A  G  P  I  Y  D  V  D  L  R  Q  E  S  H  G                139

648  TAT CTC GAC GGT ATC CCC GTG AGC TGG TAC GGC GAG CGC GAC TGG GCA AAT  698
140  Y  L  D  G  I  P  V  S  W  Y  G  E  R  D  W  A  N                156

699  CTC GGC AAG AGC CAG CAT GAG GCG CTC GCC GAC GAG CGG CAC CGC TTG CAC  749
157  L  G  K  S  Q  H  E  A  L  A  D  E  R  H  R  L  H                173

750  GCA GCG CTC CAT AAG ACG GTC TAC ATC GCG CCG CTC GGC AAG CAC AAG CTC  800
174  A  A  L  H  K  T  V  Y  I  A  P  L  G  K  H  K  L                190

801  CCC GAG GGC GGC GAA GTC CGC CGC GTA CAG AAG GTG CAG ACG GAA CAG GAA  851
191  P  E  G  G  E  V  R  R  V  Q  K  V  Q  T  E  Q  E                207

852  GTC GCC GAG GCC GCG GGG ATG CGC TAT TTC CGC ATC GCG GCG ACG GAT CAT  902
208  V  A  E  A  A  G  M  R  Y  F  R  I  A  A  T  D  H                224

903  GTC TGG CCA ACG CCG GAG AAC ATC GAC CGC TTC CTC GCG TTT TAC CGC ACG  953
225  V  W  P  T  P  E  N  I  D  R  F  L  A  F  Y  R  T                241

954  CTG CCG CAG GAT GCG TGG CTC CAT TTC CAT TGT GAA GCC GGT GTC GGC CGC 1004
242  L  P  Q  D  A  W  L  H  F  H  C  E  A  G  V  G  R                258

1005 ACG ACG GCG TTC ATG GTC ATG ACG GAT ATG CTG AAG AAC CCG TCC GTA TCG 1055
259  T  T  A  F  M  V  M  T  D  M  L  K  N  P  S  V  S                275

1056 CTC AAG GAC ATC CTC TAT CGC CAG CAC GAG ATC GGC GGC TTT TAC TAC GGG 1106
276  L  K  D  I  L  Y  R  Q  H  E  I  G  G  F  Y  Y  G                292
```

```
1107 GAG TTC CCC ATC AAG ACG AAG GAT AAA GAT AGC TGG AAG ACG AAA TAT TAT 1157
293  E   F   P   I   K   T   K   D   K   D   S   W   K   T   K   Y   Y   309

1158 AGG GAA AAG ATC GTG ATG ATC GAG CAG TTC TAC CGC TAT GTG CAG GAG AAC 1208
310  R   E   K   I   V   M   I   E   Q   F   Y   R   Y   V   Q   E   N   326

1209 CGC GCG GAT GGC TAC CAG ACG CCG TGG TCG GTC TGG CTC AAG AGC CAT CCG 1259
327  R   A   D   G   Y   Q   T   P   W   S   V   W   L   K   S   H   P   343

1260 GCG AAG GCG TAA AAGCGCAGGC GGCGGCTCGG AGTCAGGGAA ATGGCGCTGC        1311
344  A   K   A   *                                                     346

1312 CAGCACGGGA CGCGCGGCGG CGGATGCTGC GCCGGTCAGG GATGATTGAC GACAGCCAGA  1371

1372 GAAGAAAGGA TGGTTTTATG AGGTGGATCC                                   1401
```

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/55, 9/16, 1/19, 1/20, 1/21, A01H 5/00, A23K 1/00, 1/165, C12Q 1/44, 1/68, C12N 15/82 // (C12N 1/19, C12R 1:84) (C12N 1/21, C12R 1:01, 1:125, 1:19)

A3

(11) International Publication Number: WO 97/48812

(43) International Publication Date: 24 December 1997 (24.12.97)

(21) International Application Number: PCT/CA97/00414

(22) International Filing Date: 13 June 1997 (13.06.97)

(30) Priority Data:

60/019,735	14 June 1996 (14.06.96)	US
08/744,779	6 November 1996 (06.11.96)	US
08/862,531	23 May 1997 (23.05.97)	US

(71) Applicant *(for all designated States except US)*: HER MAJESTY THE QUEEN IN RIGHT OF CANADA, represented by THE DEPARTMENT OF AGRICULTURE AND AGRI-FOOD CANADA [CA/CA]; Lethbridge Research Station, P.O. Box 3000, Main Lethbridge, Alberta T1J 4B1 (CA).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: CHENG, Kuo, Joan [CA/CA]; 2015 - 6th Avenue South, Lethbridge, Alberta T1J 1C2 (CA). SELINGER, Leonard, Brent [CA/CA]; 114 Laval Court West, Lethbridge, Alberta T1K 4G3 (CA). YANKE, Lindsey, Jay [CA/CA]; 1520 - 23rd Avenue N. #18, Lethbridge, Alberta T1H 4X9 (CA). BAE, Hee, Dong [KR/KR]; Eung-Sup Shim, 114-5, Chi-dong, Paldal-gu, Suwon, Seoul 442-060 (KR). ZHOU, Luming [CA/US]; 1227 East Brickyard Road #203, Salt Lake City, UT 84106 (US). FORSBERG, Cecil, Wallace [CA/CA]; 44 Hands Drive, Guelph, Ontario N1G 3H3 (CA).

(74) Agent: McKAY-CAREY, Mary, Jane; McKay-Carey & Company, 10155 - 102nd Street, 2125 Commerce Place, Edmonton, Alberta T5J 4G8 (CA).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report: 26 March 1998 (26.03.98)

(54) Title: DNA SEQUENCES ENCODING PHYTASES OF RUMINAL MICROORGANISMS

(57) Abstract

Phytases derived from ruminal microorganisms are provided. The phytases are capable of catalyzing the release of inorganic phosphorus from phytic acid. Preferred sources of phytases include *Selenomonas, Prevotella, Treponema and Megasphaera*. A purified and isolated DNA encoding a phytase of *Selenomonas ruminantium* JY35 (ATCC 55785) is provided. Recombinant expression vectors containing DNAs encoding the phytases and host cells transformed with DNAs encoding the phytases are also provided. The phytases are useful in a wide range of applications involving the dephosphorylation of phytate, including, among other things, use in animal feed supplements.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

INTERNATIONAL SEARCH REPORT

[redacted]ial Application No
PCT/CA 97/00414

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/55 C12N9/16 C12N1/19 C12N1/20 C12N1/21
A01H5/00 A23K1/00 A23K1/165 C12Q1/44 C12Q1/68
C12N15/82 //(C12N1/19,C12R1:84),(C12N1/21,C12R1:01,1:125,1:19)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A01H A23K C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 420 358 A (GIST BROCADES NV) 3 April 1991	30
A	see page 2, line 10-12	1,2,11, 12,14, 15,18, 21,22, 24,36, 39-42, 48,51, 52,58, 69,71, 72,74
	see page 4, line 47-55 see page 7; table 1 see page 11, line 30-50; example 2 see page 19; example 9 see page 28; claims 22,23,28-31 --- -/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

6

Date of the actual completion of the international search	Date of mailing of the international search report
22 January 1998	03.02.98
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Authorized officer Macchia, G

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 449 375 A (GIST BROCADES NV ;MOGEN INT (NL)) 2 October 1991 see page 2, line 11-13	30
A	see page 3, line 45-54 see page 7, line 28-55 see page 12, line 40-58; example 10	1,27,36, 51,52, 55,58, 69,71, 72,74
X	WO 94 03072 A (PANLABS INC ;NEVALAINEN HELENA K M (AU); ALKO LTD (FI); PALOHEIMO) 17 February 1994	30
A	see page 1, paragraph 2 see page 2, paragraph 1 see page 28, paragraph 1-2 see page 53, paragraph 6	43,45, 69,71, 72,74
X	PUNJ M.L. ET AL.: "Utilization of phytin phosphorus by rumen microorganisms" THE INDIAN VETERINARY JOURNAL, vol. 46, no. 10, 1969, pages 881-886, XP002044300	30
A	see page 885	12,13
A	WO 93 16175 A (GIST BROCADES NV) 19 August 1993 see page 5, line 16 - page 6, line 7	69,71, 72,74

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

[redacted]tional application No.
PCT/CA 97/00414

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☒ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☒ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: 1-67 75-77

Isolated nucleic acid encoding a phytase of a ruminal microorganism, as in Seq.ID:1. Expression construct comprising it, transformed host cells and transgenic plants thereof. Selenomonas ruminantium JY35 (ATCC 55785). Phytase of a ruminal microorganism as in Seq.ID:2. Feed composition comprising feedstuff treated with said phytase, feed additive comprising microorganisms expressing said phytase. Method for producing said phytase. Method for producing a transgenic plant expressing said phytase. Method for improving dietary phytate utilization by an animal, via use of said phytase. Method to identify a nucleic acid encoding a phytase by means of at least part of Seq.ID:1.

2. Claims: 68-74

A method for assaying phytase activity of a microorganism.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/CA 97/00414

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0420358 A	03-04-91	AU 636673 B	06-05-93
		AU 6501190 A	28-04-91
		BG 60108 A	15-10-93
		CA 2042054 A	28-03-91
		CN 1051058 A	01-05-91
		DE 420358 T	12-10-95
		EP 0779037 A	18-06-97
		ES 2072834 T	01-08-95
		JP 4506007 T	22-10-92
		LT 1527 A,B	26-06-95
		LV 10310 A,B	20-10-94
		WO 9105053 A	18-04-91
		PL 167790 B	30-11-95
		PL 168470 B	29-02-96
		US 5436156 A	25-07-95
EP 0449375 A	02-10-91	AU 649447 B	26-05-94
		AU 7765691 A	21-10-91
		AU 632941 B	14-01-93
		AU 7776691 A	21-10-91
		CA 2056396 A	24-09-91
		EP 0449376 A	02-10-91
		IL 97645 A	18-03-97
		JP 6501838 T	03-03-94
		JP 6502296 T	17-03-94
		WO 9114782 A	03-10-91
		WO 9114772 A	03-10-91
		US 5543576 A	06-08-96
		US 5593963 A	14-01-97
WO 9403072 A	17-02-94	AU 4787993 A	03-03-94
		CA 2141431 A	17-02-94
		EP 0655890 A	07-06-95
		FI 950409 A	30-03-95
		JP 9505201 T	27-05-97
WO 9316175 A	19-08-93	AU 3628493 A	03-09-93
		EP 0626010 A	30-11-94
		FI 943707 A	10-08-94

PCT

D INTELLECTUAL PROPERTY ORGANIZATIO
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
A01N 43/04, 63/00, C12N 13/00

A1

(11) International Publication Number: WO 99/17610

(43) International Publication Date: 15 April 1999 (15.04.99)

(21) International Application Number: PCT/US98/20199

(22) International Filing Date: 24 September 1998 (24.09.98)

(30) Priority Data:
08/942,939 2 October 1997 (02.10.97) US

(71) Applicant: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 12th floor, 1111 Franklin Street, Oakland, CA 94607-5200 (US).

(72) Inventors: ROTHMAN, Stephen, S.; 98 Acacia Avenue, Berkeley, CA 94708 (US). GOLDFINE, Ira, D.; 640 Goodhill Road, Kentfield, CA 94909 (US). GERMAN, Michael, S.; 1543 33rd Avenue, San Francisco, CA 94122 (US).

(74) Agent: FRANCIS, Carol, L.; Bozicevic, Field & Francis LLP, Suite 200, 285 Hamilton Avenue, Palo Alto, CA 94301 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: PROTEIN DELIVERY BY SECRETORY GLAND EXPRESSION

(57) Abstract

Secretory gland cells, particularly pancreatic, hepatic, and salivary gland cells, are genetically altered to operatively incorporate a gene which expresses a protein which has a desired therapeutic effect on a mammalian subject. The expressed protein is secreted directly into the bloodstream to obtain therapeutic levels of the protein thereby treating the patient in need of the protein. The transformed secretory gland cells provide long term or short term therapies for diseases associated with a deficiency in a particular protein or which are amenable to treatment by overexpression of a protein.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

PROTEIN DELIVERY BY SECRETORY GLAND EXPRESSION

Field of the Invention

This invention relates generally to the field of genetic transformation of cells *in vivo*, more particularly to *in vivo* transformation of secretory gland cells by introduction of the transforming nucleic acid into a secretory gland via a secretory gland duct.

Background of the Invention

The ability to replace defective or absent genes has attracted wide attention as a method to treat a variety of human diseases (Crystal 1995 Science 270:404), Lever et al. 1995 Gene Therapy. Pearson Professional. New York p. 1-91; Friedmann 1996 Nature Med. 2:144). Although originally intended as a means of correcting inherited disorders in certain populations of somatic cells, gene-based therapy can be a useful means to supply exogenous gene products to the circulatory system for the treatment of a wide range of systemic disorders that involve deficiencies in circulating proteins, such as hormones, growth factors, and clotting proteins (Lever et al. 1995 supra; Buckel 1996 TiPS 17:450), as well as a means of administering other polypeptide drugs. The success of this application depends upon developing effective methods to both manufacture the desired protein *in vivo* and then secrete it into blood (Crystal 1995 supra; Lever et al. 1995 supra).

Currently, DNA-based therapy (*i.e.*, gene therapy) is carried out in a variety of ways but involves two general protocols. In the first method, referred to as *ex vivo* gene therapy, cells are extracted from an individual and subjected to genetic manipulation. After genetic material has been properly inserted into the cells, the cells are implanted back into the individual from which they were removed. Persistent, *in vivo* expression of the newly implanted genetic material after transplantation of the transformed cells has been successful (see Morgan et al., *Science* **237**:1476 (1987); and Gerrard et al., *Nat. Genet.* 3:180 (1993)). In the second approach to DNA-based therapy, referred to as *in vivo* gene therapy, cells within a living organism are transformed *in situ* with exogenous genetic material.

Several different methods for transforming cells can be used in accordance with either the *ex vivo* or *in vivo* transfection procedures. For example, various mechanical methods can be used to deliver the genetic material, including the use of fusogenic lipid vesicles (liposomes incorporating cationic lipids such as lipofection; see Felgner et al., *Proc. Natl. Acad. Sci. U.S.A.* **84**:7413-7417 (1987)); direct injection of DNA (Wolff, et al., *Science* (1990) **247**:1465-1468); and pneumatic delivery of DNA-coated gold particles with a device referred to as the gene gun (Yang et al., *Proc. Natl. Acad. Sci. U.S.A.* 1990; 87:1568-9572). Morsy et al. reviews several of the different

techniques useful in transformation of cells *ex vivo* or *in vivo* and provides citations of numerous publications in each area (Morsy et al., *JAMA* **270**:2338-2345 (1993)).

One method of particular interest for delivery of genetic material involves use of recombinant viruses to infect cells *in vivo* or *ex vivo*. In these methods, a virus containing the desired genetic material is allowed to infect target cells within the subject. Upon infection, the virus injects its genetic material into the target cells. The genetic material is then expressed within the target cell, providing for expression of the desired genetic material. However, it would be preferable to avoid introduction of the desired genetic material by viral infection for a number of reasons. For example, viral infection results in delivery of viral DNA in addition to the desired genetic material, which may in turn result in undesirable cellular effects such as, adverse immune reactions, productive viral replication, and adverse integration events.

There is a need in the field for a method for delivery of genetic material into a cell *in vivo* to provide for expression of the introduced polynucleotide and secretion of the gene product it encodes into the bloodstream. The present invention addresses this problem.

Summary of the Invention

Secretory gland cells are genetically altered to operatively incorporate a gene which is expressed by the genetically altered secretory gland cell to produce a polypeptide which is subsequently secreted into the bloodstream. Specifically, the invention involves introduction of a nucleotide of interest into a secretory gland via the duct system (*e.g.*, by retrograde ductal administration) to transform a secretory gland cell.

In one embodiment the invention features genetic alteration of cells of two secretory glands (*e.g.*, the liver and the pancreas).

In another embodiment, the invention features transformation of pancreatic cells with insulin-encoding nucleic acid to provide for expression and secretion of insulin at levels sufficient to maintain a substantially euglycemic state in a subject having a diabetic syndrome.

A primary object is to provide a non-invasive method of protein delivery (*i.e.*, the method involves introduction of the nucleic acid of interest from outside the body (*i.e.*, from the duct system of particular glands) wherein cells of a secretory gland, preferably the pancreas, salivary gland, or liver of a mammal are genetically modified to express a biologically active and therapeutically useful polypeptide, which polypeptide is secreted into the circulatory system of the individual.

Another object is to produce genetically transformed secretory gland cells which cells have incorporated into their genome genetic material which, when expressed, produces a biologically active and therapeutically useful protein which is secreted into the circulatory system.

An advantage of the present invention is that both long and short term therapy can be provided for diseases wherein individuals are suffering from the disease due to a deficiency in a particular protein, or by supplying an exogenous protein having a desired activity (e.g. antimicrobial activity).

These and other objects, advantages and features of the present invention will become apparent to those persons skilled in the art upon reading the details of the methodology and compositions as more fully set forth below.

Brief Description of the Drawing

Fig. 1 is a map of the pFGH construct, which contains the human growth hormone (hGH) genomic sequence.

Fig. 2 is a map of the pFGH.CMV construct, which contains the hGH genomic sequence operably linked to the CMV promoter.

Fig. 3 is a map of the pFGH.chymo construct, which contains the hGH genomic sequence operably linked to the chymotrypsin B promoter.

Fig. 4 is a graph showing the levels of tissue expression of hGH expression in the pancreas of rats after retrograde injection with either a control containing no DNA or a test sample containing a hGH construct.

Fig. 5 is a graph showing the serum levels of hGH in rats after retrograde pancreatic injection with either a control containing no DNA or a test sample containing a human growth hormone construct.

Fig. 6 is a graph showing the correlation between pancreatic tissue expression and serum levels of hGH.

Fig. 7 is a graph showing tissue expression of hGH following transformation of salivary gland cells by intraductal injection.

Fig. 8 is a graph showing regulation of plasma levels of hGH expressed from transformed salivary gland cells.

Fig. 9 is a graph showing regulation of plasma levels of hGH expressed from transformed pancreatic cells.

Fig. 10 is a map of the pBAT16.hInsG1.M2 construct, which contains DNA encoding an altered form of human insulin.

Fig. 11 is a graph showing the glucose response in streptozotocin-treated rats having pancreatic cells transformed with either human insulin (open bars) or green fluorescent protein (GFP; striped bars).

Fig. 12, is a graph showing the blood glucose levels in control rats (mock-treated; closed squares), streptozotocin-treated rats (open squares), and streptozotocin rats treated by transformation of pancreatic cells with DNA encoding human insulin (closed circles).

Fig. 13 is a graph showing expression of hGH in the plasma of control rats (no DNA) and of rats in which hGH-encoding DNA was introduced into the liver by intraductal injection.

Fig. 14 is a graph showing the relative amounts of hGH in the pancreatic tissue of rats that received either pFGH (control), pFGH.chymo, pFGH.RSV, pFGH.RSV, or pFGH.CMV by intraductal administration to the pancreas.

Fig. 15 is a graph showing the relative amounts of hGH in the pancreatic tissue of rats that received either no DNA (mock-transformed), pFGH.CMV, pFGH.CMV premixed with lipofectin, or pFGH.CMV premixed with adenovirus.

Fig. 16 is a graph showing the relative levels of plasma hGH in rats that received either pFGH (control), pFGH.chymo, pFGH.RSV, pFGH.RSV, or pFGH.CMV by intraductal administration to the pancreas.

Fig. 17 is a graph showing the relative amounts of plasma hGH in rats that received either no DNA (mock-transformed), pFGH.CMV, pFGH.CMV premixed with lipofectin, or pFGH.CMV premixed with adenovirus by intraductal administration to the pancreas..

Fig. 18 is a graph showing the relative levels of plasma hGH in rats that received no DNA (control), received hGH-encoding DNA via intraductal delivery to the liver, received hGH-encoding DNA via intraductal delivery to the pancreas, or received hGH-encoding DNA via intraductal delivery to both the liver and pancreas.

Fig. 19 is a graph showing the relative levels of hGH expression in pancreas tissue following administration of DNA to both pancreas and liver or to pancreas alone. The graph shows tissue levels of hGH after administration of a control (no DNA) to both pancreas and liver (left-most bar); administration of pFGH.CMV to both pancreas and liver (center bar); and pFGH.CMV to pancreas alone (right-most bar). Adenovirus was admixed with the construct as an adjuvant.

Fig. 20 is graph showing the relative levels of hGH expression in salivary gland tissue in rats that received either no DNA (control rats), pFGH.CMV, pFGH.CMV premixed with lipofectin, or pFGH.CMV premixed with adenovirus.

Fig. 21 is a graph showing stimulation of hGH secretion into the plasma of rats that received hGH-encoding DNA by intraductal injection into both the pancreas and liver.

Fig. 22 is a graph showing the blood glucose levels of streptozotocin-treated rats (diabetic) that received either no DNA (open squares) or received human insulin-encoding DNA by intraductal injection into the pancreas (closed squares) over a three day period.

Fig. 23 is a graph showing the plasma insulin levels of streptozotocin-treated rats (diabetic) that received either no DNA (open squares) or received human insulin-encoding DNA by intraductal injection into the pancreas (closed squares)over a three day period.

Fig. 24 is a graph showing the blood glucose levels (over a six day period) of streptozotocin-treated rats (diabetic) that received either no DNA (open squares) or received human insulin-encoding DNA by intraductal injection into the pancreas (closed squares).

Description of the Preferred Embodiments

Before the present method of genetically transforming secretory gland cells and methods for protein delivery are described, it is to be understood that this invention is not limited to the particular methodology, protocols, cell lines, secretory glands, vectors and reagents described as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a", "and", and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a secretory gland cell" includes a plurality of such cells and reference to "the transformation vector" includes reference to one or more transformation vectors and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood to one of ordinary skill in the art to which this invention belongs. Although any methods, devices and materials similar or equivalent to those described herein can be used in the practice or testing of the invention, the preferred methods, devices and materials are now described.

All publications mentioned herein are incorporated herein by reference for the purpose of describing and disclosing the cell lines, vectors, and methodologies which are described in the publications which might be used in connection with the presently described invention. The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such a disclosure by virtue of prior invention.

Definitions

By "secretory gland" is meant an aggregation of cells specialized to secrete or excrete materials not related to their ordinary metabolic needs. Secretory glands include salivary glands,

pancreas, mammary glands, thyroid gland, thymus gland, pituitary gland, liver, and other glands well known in the art.

By "exocrine gland" is meant a ducted gland or portion of a ducted gland that releases its products externally relative to the body, *e.g.*, either into the internal cavities such as the ocular and nasal cavities, the lumen of the gastrointestinal tract, or onto the surface of the body.

By "salivary gland" is meant a gland of the oral cavity which secretes saliva, including the glandulae salivariae majores of the oral cavity (the parotid, sublingual, and submandibular glands) and the glandulae salivariae minores of the tongue, lips, cheeks, and palate (labial, buccal, molar, palatine, lingual, and anterior lingual glands).

By "pancreas" is meant a large, elongated, racemose gland situated transversely behind the stomach, between the spleen and the duodenum. The pancreas is composed of an endocrine portion (the pars endocrina) and an exocrine portion (the pars exocrina). The pars endocrina, which contains the islets of Langerhans, produces and secretes proteins, including insulin, directly into the bloodstream. The pars exocrina contains secretory units and produces and secretes a pancreatic juice, which contains enzymes essential to protein digestion, into the duodenum.

By "retrograde ductal injection" is meant the administration of a liquid or other material into the fluid contents of the duct system of an exocrine gland in a direction opposite to the normal flow of that fluid, either at the external orifice of the duct system or through its wall. "Retrograde ductal injection" can be a single, discontinuous administration or continuous administration (*i.e.*, perfusion).

By "transformation" is meant a genetic change induced in a cell following incorporation of new DNA (*i.e.*, DNA exogenous to the cell). Where the cell is a mammalian cell, the genetic change may be achieved by introduction of the DNA into the genome of the cell.

By "transfection" is meant the transformation of a cell with DNA from a virus.

By "transformed cell" is meant a cell into which (or, where the introduced DNA is incorporated into the genome, into an ancestor of which) has been introduced, by means of recombinant DNA techniques, a DNA molecule encoding a protein of interest.

By "nucleic acid of interest" is meant any DNA or RNA molecule which encodes a polypeptide or other molecule which is desirable for administration to a mammalian subject for expression of the product encoded by the nucleic acid of interest and delivery of the encoded product into the blood stream of the mammalian subject. The nucleic acid is generally operatively linked to other sequences which are needed for its expression such as a promoter. The term "DNA of interest" is used as shorthand herein to refer to the nucleic acid of interest.

By "construct" is meant a nucleic acid molecule which contains the nucleic acid of interest (*e.g.*, the DNA of interest), generally operably linked to a promoter for expression of the polypeptide

encoded by the nucleic acid of interest. "Constructs" as used herein is generally meant to refer a nucleic acid molecule that facilitates expression of a polypeptide encoded by the nucleic acid to be introduced into a secretory gland cell.

By "vector" is meant any compound, biological or chemical, that facilitates transformation of a secretory gland cell with a DNA of interest. Exemplary biological vectors include viruses, particularly attenuated and/or replication-deficient viruses. Exemplary chemical vectors include lipid complexes and naked DNA constructs.

By "naked DNA" or "naked nucleic acid" or DNA sequence and the like is meant a nucleic acid molecule that is not contained within a viral particle, bacterial cell or other encapsulating means that facilitates delivery of nucleic acid into the cytoplasm of the target cell. Naked nucleic acid can be associated with means for facilitating delivery of the nucleic acid to the site of the target cell (*e.g.*, means that facilitate travel into the target cell of the nucleic acid through the alimentary canal, protect the nucleic acid from stomach acid, and/or serve to penetrate intestinal mucus) and/or to the surface of the target epithelial cell.

By "promoter" is meant a minimal DNA sequence sufficient to direct transcription. "Promoter" is also meant to encompass those promoter elements sufficient for promoter-dependent gene expression controllable for cell-type specific, tissue-specific or inducible by external signals or agents; such elements may be located in the 5' or 3' regions of the native gene.

By "secretory gland specific promoter" is meant a promoter which directs expression of an operably linked DNA sequence when bound by transcriptional activator proteins, or other regulators of transcription, which are unique to a specific type of secretory gland cell. For example, by "salivary gland specific promoter" is meant a secretory gland specific promoter which directs expression in a salivary gland cell. A salivary amylase promoter is an example of a salivary gland specific promoter. By "pancreas specific promoter" is meant a secretory gland specific promoter which directs expression in a pancreatic cell. Examples of pancreas specific promoters include a pancreatic amylase promoter and an insulin promoter.

By "operably linked" is meant that a DNA sequence and a regulatory sequence(s) are connected in such a way as to permit gene expression when the appropriate molecules (*e.g.*, transcriptional activator proteins) are bound to the regulatory sequence(s).

By "operatively inserted" is meant that the DNA of interest introduced into the cell is positioned adjacent a DNA sequence which directs transcription and translation of the introduced DNA (*i.e.*, facilitates the production of, *e.g.*, a polypeptide encoded by a DNA of interest).

By "mammalian subject" or "mammalian patient" is meant any mammal to which intravenous protein delivery is desired, including human, bovine, equine, canine, and feline subjects.

By "euglycemia" or "euglycemic state" is meant a state associated with a level of blood glucose that is normal or nearly normal, particularly relative to the levels of blood glucose in a subject having a disease or condition associated with hyperglycemia. In humans, euglycemia correlates with blood glucose levels in the range of 70 mg/dl to 130 mg/dl.

The terms "synergistic," "synergistic effect," and the like are used herein to describe improved effects (e.g., an increase in tissue expression levels in one or more secretory glands, an increased responsiveness to hormonal stimulation to elicit secretion of a polypeptide of interest, or a decrease in an undesirable phenotype) by combining one or more aspects of the invention (*e.g.*, by transformation of more than one secretory gland in a single subject, or by transformation of a secretory gland(s) with multiple constructs encoding the same or different polypeptides).

Overview of the Invention

The present invention features methods for genetically altering a secretory gland cell (*i.e.*, secretory gland cell transformation) and methods of delivering a protein using the methods of genetically altering secretory gland cells. More specifically the invention features methods for delivery of a protein or other product encoded by a nucleic acid sequence of interest to a mammalian subject by expression of a DNA of interest in cells within a secretory gland of a mammalian patient (*i.e.*, by *in vivo* gene therapy). Preferably, the transformed secretory gland cells expressing the protein encoded by the DNA of interest secrete a therapeutically effective amount of the protein into the bloodstream of the mammalian patient. Preferably, the secretory gland into which the DNA of interest is introduced and expressed is the pancreas, a salivary gland, or the liver. In short, the invention features a delivery system that involves introduction of a nucleic acid sequence encoding a product of interest (*e.g.*, a protein) into a secretory gland cell (*e.g.*, a salivary gland cell, hepatocyte, or pancreatic cell, particularly exocrine cells of salivary gland, liver, or pancreas), expression of the encoded protein, and delivery of the protein into the blood stream by secretion of the protein by the transformed secretory gland cell.

The present invention preferably uses either naked DNA or DNA premixed with adjuvants (*e.g.*, lipofectin or viral particles). It is not necessary to incorporate the DNA into viral particles in order to achieve transformation of secretory gland cells and provide expression of the polypeptide of interest at physiologic/therapeutic levels in the bloodstream.

An important feature of the invention is the use of exocrine cells of glands of the gastrointestinal tract (*i.e.*, pancreas, liver, salivary gland) to produce and secrete therapeutic proteins into blood. While it is well understood that exocrine cells secrete into the lumen of the glands' ducts (*i.e.* in an exocrine direction), with the exception of the liver (*i.e.*, the hepatocytes secrete cellular products in both directions, *e.g.* blood proteins into blood and bile salts into the intestinal lumen), it

is not widely appreciated that exocrine cells can also secrete significant amounts of protein into the systemic circulation. For example, exocrine proteins such as α-amylase (salivary glands), pepsinogen (gastric glands), various digestive enzymes from the exocrine pancreas, salivary gland kallikreins and nerve growth factor (Liebow, 1988 Pancreas 3:343-351) are normal constituents of blood. In the pancreas, substantial quantities of digestive enzymes are released into the circulation (Saito et al., 1973 Jpn. J. Physiol. 23:477-95; Isenman et al. 1997 Proc. Natl. Acad. Sci (USA) 74:4068-4072; Papp et al.1980 Acta Physiol. Acad. Sci. Hung. 56:401-410; Geokas et al., 1980 Am. J. Physiol. 238:238-246; Miyasaka et al. 1981 Am. J. Physiol. 241:170-175; Grendell et al. 1982 Am. J. Physiol. 243:54-59). Endocrine secretion can be greatly enhanced by common secretory stimulants (Saito et al., *supra*; Isenman et al. *supra*; Miyasaka et al. *supra*; Grendell et al. *supra*). As much as 20-25% of the total secreted product can be released into blood as a consequence of stimulation (Grendell et al. *supra*). The present invention takes advantage of the discovery that exocrine gland cells can be transformed with a desired DNA sequence and secrete the encoded polypeptide into the bloodstream rather than only or primarily into the gastrointestinal tract.

In addition to the advantages described above, the invention also permits access to the cells of secretory glands without invasive procedures. For example, it is possible to cannulate either the collecting duct of a major salivary gland through its orifice in the mouth, or the common bile or pancreatic duct by means of endoscopic retrograde cholangiopancreatography (ERCP). These are common diagnostic procedures performed on awake patients. The non-invasive methods of the invention allow delivery of the DNA of interest in a safe manner that substantially avoids the inflammatory and immunological responses associated with other means of DNA delivery.

The invention also takes advantage of the protein-producing capacity of secretory gland cells. This advantage is particularly useful for the production of hormones such as hGH and insulin, which have short half-lives in blood and are cleared quickly. The cells of the exocrine glands are the body's major protein synthesizing and secreting systems. For example, the human exocrine pancreas manufactures and secretes approximately 20 g of protein daily. According to the present invention, even a small proportion of protein synthesized by secretory glands provides enough secreted product to provide therapeutic protein levels for the treatment of most diseases of circulating proteins.

The invention will now be described in further detail.

Constructs

Any nucleic acid construct having a eukaryotic promoter operably linked to a DNA of interest can be used in the invention. The constructs containing the DNA sequence (or the corresponding RNA sequence) which may be used in accordance with the invention may be any eukaryotic expression construct containing the DNA or the RNA sequence of interest. For example,

a plasmid or viral construct (*e.g.* adenovirus) can be cleaved to provide linear DNA having ligatable termini. These termini are bound to exogenous DNA having complementary, like ligatable termini to provide a biologically functional recombinant DNA molecule having an intact replicon and a desired phenotypic property. Preferably the construct is capable of replication in both eukaryotic and prokaryotic hosts, which constructs are known in the art and are commercially available.

The exogenous (*i.e.*, donor) DNA used in the invention is obtained from suitable cells, and the constructs prepared using techniques well known in the art. Likewise, techniques for obtaining expression of exogenous DNA or RNA sequences in a genetically altered host cell are known in the art (see, for example, Kormal et al., *Proc. Natl. Acad. Sci. USA*, **84**:2150-2154, 1987; Sambrook et al. Molecular Cloning: a Laboratory Manual, 2nd Ed., 1989, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY; each of which are hereby incorporated by reference with respect to methods and compositions for eukaryotic expression of a DNA of interest).

Preferably, the DNA construct contains a promoter to facilitate expression of the DNA of interest within a secretory gland cell. Preferably the promoter is a strong, eukaryotic promoter such as a promoter from cytomegalovirus (CMV), mouse mammary tumor virus (MMTV), Rous sarcoma virus (RSV), or adenovirus. More specifically, exemplary promoters include the promoter from the immediate early gene of human CMV (Boshart et al., *Cell* **41**:521-530, 1985) and the promoter from the long terminal repeat (LTR) of RSV (Gorman et al., *Proc. Natl. Acad. Sci. USA* **79**:6777-6781, 1982). Of these two promoters, the CMV promoter is preferred as it provides for higher levels of expression than the RSV promoter.

Alternatively, the promoter used may be a tissue-specific promoter. For example, where the secretory gland is the pancreas, the promoter used in the vector is preferably a pancreas specific promoter, *e.g.*, an insulin promoter or a pancreas α-amylase promoter; where the secretory gland is a salivary gland, the tissue-specific promoter may be a salivary α-amylase promoter or mumps viral gene promoter. Both pancreatic and salivary α-amylase genes have been identified and characterized in both mice and humans (see, for example, Jones et al., *Nucleic Acids Res.*, **17**:6613-6623; Pittet et al., *J. Mol. Biol.*, **182**:359-365, 1985; Hagenbuchle et al., *J. Mol. Biol.*, **185**:285-293, 1985; Schibler et al., *Oxf. Surv. Eukaryot. Genes*, **3**:210-234, 1986; and Sierra et al., *Mol. Cell. Biol.*, 6:4067-4076, 1986 for murine pancreatic and salivary α-amylase genes and promoters; Samuelson et al., *Nucleic Acids Res.*, **16**:8261-8276, 1988; Groot et al., *Genomics*, **5**:29-42, 1989; and Tomita et al., *Gene*, **76**:11-18, 1989 for human pancreatic and salivary α-amylase genes and their promoters; Ting et al., *Genes Dev.* **6**:1457-65, 1992 for human salivary α-amylase AMY1C promoter sequences).

The constructs of the invention may also include sequences in addition to promoters which enhance secretory gland specific expression. For example, where pancreas specific expression of the

DNA of interest is desired, the construct may include a PTF-1 recognition sequence (Cockell et al., *Mol. Cell. Biol.*, 9:2464-2476, 1989). Sequences which enhance salivary gland specific expression are also well known in the art (see, for example, Robins et al., *Genetica* 86:191-201, 1992).

Other components such as a marker (*e.g.*, an antibiotic resistance gene (such as an ampicillin resistance gene) or β-galactosidase) to aid in selection of cells containing and/or expressing the construct, an origin of replication for stable replication of the construct in a bacterial cell (preferably, a high copy number origin of replication), a nuclear localization signal, or other elements which facilitate production of the DNA construct, the protein encoded thereby, or both.

For eukaryotic expression, the construct should contain at a minimum a eukaryotic promoter operably linked to a DNA of interest, which is in turn operably linked to a polyadenylation sequence. The polyadenylation signal sequence may be selected from any of a variety of polyadenylation signal sequences known in the art. Preferably, the polyadenylation signal sequence is the SV40 early polyadenylation signal sequence. The construct may also include one or more introns, which can increase levels of expression of the DNA of interest, particularly where the DNA of interest is a cDNA (*e.g.*, contains no introns of the naturally-occurring sequence). Any of a variety of introns known in the art may be used (e.g, the human β-globin intron, which is inserted in the construct at a position 5' to the DNA of interest).

The DNA of interest may be inserted into a construct so that the therapeutic protein is expressed as a fusion protein (*e.g.*, a fusion protein having β-galactosidase or a portion thereof at the N-terminus and the therapeutic protein at the C-terminal portion). Production of a fusion protein can facilitate identification of transformed cells expressing the protein (*e.g.*, by enzyme-linked immunosorbent assay (ELISA) using an antibody which binds to the fusion protein).

The Nucleic Acid (DNA) of Interest

The DNA of interest can be any DNA encoding any protein for which intravenous therapy is desirable. For example, intravenous protein therapy is appropriate in treating a mammalian subject having an inherited or acquired disease associated with a specific protein deficiency (*e.g.*, diabetes, hemophilia, anemia, severe combined immunodeficiency). Such protein deficient states are amenable to treatment by replacement therapy, *i.e.*, expression of a protein to restore the normal bloodstream levels of the protein to at least normal levels.

Alternatively, the DNA of interest may encode a polypeptide that is either normally present in a healthy mammalian subject or which is foreign to the mammalian subject, and which polypeptide is effective in treatment of a condition by expression or over-expression of the polypeptide. For example, the DNA of interest can encode antimicrobial, antiparasitic, antifungal, or antiviral polypeptides for treatment of a mammalian subject having a viral (*e.g.*, human

immunodeficiency virus (HIV), Epstein-Barr virus (EBV), herpes simplex virus (HSV), bacterial, fungal, and/or parasitic infection, particularly where the infection is chronic, *i.e.*, persisting over a relatively long period of time. The methods of the invention may also be used to enhance expression of a protein present in a normal mammal, or to express a protein not normally present in a normal mammal, in order to achieve a desired effect (*e.g.*, to enhance a normal metabolic process). For example, a secretory gland of a dairy cow may be transformed with DNA encoding bovine growth hormone (BGH) in order to enhance levels of BGH in the bloodstream and enhance milk production.

The DNA of interest is preferably obtained from a source of the same species as the mammalian subject to be treated (*e.g.* human to human), but this is not an absolute requirement. DNA obtained from a species different from the mammalian subject can also be used, particularly where the amino acid sequences of the proteins are highly conserved and the xenogeneic protein is not highly immunogenic so as to elicit a significant, undesirable antibody response against the protein in the mammalian host.

Exemplary, preferred DNAs of interest include DNA encoding insulin, growth factors (*e.g.*, growth hormone, insulin-like growth factor-1 (IGF-I), platelet-derived growth factor (PDGF), epidermal growth factor (EGF), acidic fibroblast growth factor, basic fibroblast growth factor, or transforming growth factor β), cytokines (*e.g.*, interferon (INF) (*e.g.*, INF-α2b, INF-α2a, INF-αN1, INF-β1b, INF-γ), interleukin (e.g, IL-2, IL-8), or tumor necrosis factor (TNF) (e.g, TNF--α, TNF-β)), clotting factors (*e.g.*, clotting factor VIII), hormones (e.g, GP-1), antimicrobial polypeptides (*e.g.*, antibacterial, antifungal, antiviral, and/or antiparasitic polypeptides), enzymes (*e.g.*, adenosine deaminase), filgastim (Neupogen), hemoglobin, erythropoietin, insulinotropin, imiglucerase, sarbramostim, antigens, tissue plasminogen activator (tPA), urokinase, streptokinase, endothelian, soluble CD4, and antibodies and/or antigen-binding fragments (e.g, FAbs) thereof (*e.g.*, orthoclone OKT-e (anti-CD3), GPIIb/IIa monoclonal antibody). Preferably, the mammalian subject is a human subject and the DNA expressed encodes a human protein.

Table 1 provides a list of exemplary proteins and protein classes which can be delivered to the bloodstream of a mammalian subject via the method of secretory gland cell transformation of the invention.

TABLE 1: Exemplary Proteins and Protein Classes for Expression in and Secretion by Secretory Gland Cells

SPECIFIC EXEMPLARY PROTEINS	
insulin	interferon-α2B
human growth hormone (hGH)	transforming growth factor (TGF)
erythropoietin (EPO)	ciliary neurite transforming factor (CNTF)
clotting factor VIII	insulin-like growth factor-1 (IGF-1)
bovine growth hormone (BGH)	granulocyte macrophage colony stimulating factor (GM-CSF)
platelet derived growth factor (PDGF)	interferon-α2A
clotting factor VIII	brain-derived neurite factor (BDNF)
thrombopoietin (TPO)	insulintropin
IL-1	tissue plasminogen activator (tPA)
IL-2	urokinase
IL-1 RA	streptokinase
superoxide dismutase (SOD)	adenosine deamidase
catalase	calcitonin
fibroblast growth factor (FGF) (acidic or basic)	arginase
neurite growth factor (NGF)	phenylalanine ammonia lyase
granulocyte colony stimulating factor (G-CSF)	γ-interferon
L-asparaginase	pepsin
uricase	trypsin
chymotrypsin	elastase
carboxypeptidase	lactase
sucrase	intrinsic factor
calcitonin	parathyroid hormone(PTH)-like hormone
Ob gene product	cholecystokinin (CCK)
glucagon	insulinotrophic hormone

EXEMPLARY CLASSES OF PROTEINS	
enzymes (*e.g.*, proteases, phospholipases, etc.)	pituitary hormones
protease inhibitors	growth factors
cytokines	somatomedians
chemokines	immunoglobulins
gonadotrophins	interleukins
chemotactins	interferons
lipid-binding proteins	

Numerous proteins that are desirable for intravenous protein therapy are well known in the art and the DNA encoding these proteins has been isolated. For example, the sequence of the DNAs encoding insulin, human growth hormone, intrinsic factor, clotting factor VIII, and erythropoietin are available from Genbank and/or have been described in the scientific literature (*e.g.*, human clotting factor VIII gene: Gitschier et al., *Nature* **312**:326-330, 1984; Wood et al., *Nature* **312**:330-337, 1984; human intrinsic factor: Hewitt et al., *Genomics* **10**:432-440, 1991). Moreover, proteins commonly used in treatments can be used in the procedures of the present invention. Such proteins are disclosed in, for example, the Physicians' Desk Reference (1994 Physicians' Desk Reference,

48th Ed., Medical Economics Data Production Co., Montvale, NJ; incorporated by reference) and can be dosed using methods described in Harrison's Principles of Internal Medicine and/or the AMA "Drug Evaluations Annual" 1993, all incorporated by reference.

Where the DNA encoding a protein of interest has not been isolated, this can be accomplished by various, standard protocols well known to those of skill in the art (see, for example, Sambrook et al., *ibid*; Suggs et al., *Proc. Natl. Acad. Sci. USA* **78**:6613-6617, 1981; USPN 4,394,443; each of which are incorporated herein by reference with respect to identification and isolation of DNA encoding a protein of interest). For example, genomic or cDNA clones encoding a specific protein can be isolated from genomic or cDNA libraries using hybridization probes designed on the basis of the nucleotide or amino acid sequences for the desired gene. The probes can be constructed by chemical synthesis or by polymerase chain reaction (PCR) using primers based upon sequence data to amplify DNA fragments from pools or libraries (USPNs 4,683,195 and 4,683,202). Nucleotide substitutions, deletions, additions, and the like can also be incorporated into the polynucleotides, so long as the ability of the polynucleotide to hybridize is not substantially disrupted. (Sambrook et al. *ibid*). The clones may be expressed or the DNA of interest can be excised or synthesized for use in other constructs. If desired, the DNA of interest can be sequenced using methods well known in the art.

It may also be desirable to produce altered forms of the therapeutic proteins that are, for example, protease resistant or have enhanced activity relative to the wild-type protein. For example, where the therapeutic protein is a hormone, it may be desirable to alter the protein's ability to form dimers or multimeric complexes. For example, insulin may be modified so as to prevent its dimerization has a more rapid onset of action relative to wild-type, dimerized insulin.

Vectors for Delivery of the DNA of Interest to the Secretory Gland Cell

The vectors for delivery of the DNA of interest can be either viral or non-viral, or may be composed of naked DNA admixed with an adjuvant such as viral particles (e.g. adenovirus) or cationic lipids or liposomes. An "adjuvant" is a substance that does not by itself produce the desired effect, but acts to enhance or otherwise improve the action of the active compound. The precise vector and vector formulation used will depend upon several factors such as the secretory gland targeted for gene transfer.

Non-viral vectors

The DNA of interest may be administered using a non-viral vector. "Non-viral vector" as used herein is meant to include naked DNA, chemical formulations containing naked DNA (e.g. a formulation of DNA and cationic compounds (*e.g.*, dextran sulfate)), and naked DNA mixed with an adjuvant such as a viral particle (*i.e.*, the DNA of interest is not contained within the viral particle,

but the transforming formulation is composed of both naked DNA and viral particles (*e.g.*, adenovirus particles) (see, *e.g.*, Curiel et al. 1992 Am. J. Respir. Cell Mol. Biol. 6:247-52). Thus "non-viral vector" can include vectors composed of DNA plus viral particles where the viral particles do not contain the DNA of interest within the viral genome.

In one preferred embodiment, the formulation comprises viral particles which are mixed with the naked DNA construct prior to administration. Preferably, about 10^8 to 10^{10} viral particles (preferably about 1×10^{10} to 5×10^{10}, more preferably about 3×10^{10} particles) are mixed with the naked DNA construct (about 5 μg to 50 μg DNA, more preferably about 8 μg to 25 μg DNA) in a total volume of about 100 μl. Preferably the viral particles are adenovirus particles (Curiel et al., 1992 supra).

Alternatively or in addition, the DNA of interest can be complexed with polycationic substances such as poly-L-lysine or DEAC-dextran, targeting ligands, and/or DNA binding proteins (e.g., histones). DNA- or RNA-liposome complex formulations comprise a mixture of lipids which bind to genetic material (DNA or RNA) and facilitate delivery of the nucleic acid into the cell. Liposomes which can be used in accordance with the invention include DOPE (dioleyl phosphatidyl ethanol amine), CUDMEDA (N-(5-cholestrum-3-β-ol 3-urethanyl)-N',N'-dimethylethylene diamine).

For example, the naked DNA can be administered in a solution containing Lipofectin™ (LTI/BRL) at a concentrations ranging from about 2.5 % to 15% volume:volume, preferably about 6% to 12% volume:volume. Preferred methods and compositions for formulation of DNA for delivery according to the method of the invention are described in USPN 5,527,928, incorporated herein by reference.

The DNA of interest can also be administered as a chemical formulation of DNA or RNA coupled to a carrier molecule (*e.g.*, an antibody or a receptor ligand) which facilitates delivery to host cells for the purpose of altering the biological properties of the host cells. By the term "chemical formulations" is meant modifications of nucleic acids to allow coupling of the nucleic acid compounds to a carrier molecule such as a protein or lipid, or derivative thereof. Exemplary protein carrier molecules include antibodies specific to the cells of a targeted secretory gland or receptor ligands, *i.e.*, molecules capable of interacting with receptors associated with a cell of a targeted secretory gland.

Viral vectors

In general, viral vectors used in accordance with the invention are composed of a viral particle derived from a naturally-occurring virus which has been genetically altered to render the virus replication-defective and to express a recombinant gene of interest in accordance with the invention. Once the virus delivers its genetic material to a cell, it does not generate additional

infectious virus but does introduce exogenous recombinant genes into the cell, preferably into the genome of the cell.

Numerous viral vectors are well known in the art, including, for example, retrovirus, adenovirus, adeno-associated virus, herpes simplex virus (HSV), cytomegalovirus (CMV), vaccinia and poliovirus vectors. Retroviral vectors are less preferred since retroviruses require replicating cells and secretory glands are composed of mostly slowly replicating and/or terminally differentiated cells. Adenovirus is a preferred viral vector since this virus efficiently infects slowly replicating and/or terminally differentiated cells. The viral vector may be selected according to its preferential infection of the targeted secretory gland (*e.g.*, where the secretory gland is a salivary gland, the viral vector may be derived from an attenuated (*i.e.*, does not cause significant pathology or morbidity in the infected host, e.g, the virus is nonpathogenic or causes only minor disease symptoms) and/or replication-deficient mumps virus or other attenuated and/or replication-deficient virus which is substantially specific for salivary gland cells).

Where a replication-deficient virus is used as the viral vector, the production of infective virus particles containing either DNA or RNA corresponding to the DNA of interest can be produced by introducing the viral construct into a recombinant cell line which provides the missing components essential for viral replication. Preferably, transformation of the recombinant cell line with the recombinant viral vector will not result in production of replication-competent viruses, *e.g.*, by homologous recombination of the viral sequences of the recombinant cell line into the introduced viral vector. Methods for production of replication-deficient viral particles containing a nucleic acid of interest are well known in the art and are described in, for example, Rosenfeld et al., *Science* **252**:431-434, 1991 and Rosenfeld et al., *Cell* **68**:143-155, 1992 (adenovirus); USPN 5,139,941 (adeno-associated virus); USPN 4,861,719 (retrovirus); and USPN 5,356,806 (vaccinia virus). Methods and materials for manipulation of the mumps virus genome, characterization of mumps virus genes responsible for viral fusion and viral replication, and the structure and sequence of the mumps viral genome are described in Tanabayashi et al., *J. Virol.* **67**:2928-2931, 1993; Takeuchi et al., *Archiv. Virol.*, **128**:177-183, 1993; Tanabayashi et al., *Virol.* **187**:801-804, 1992; Kawano et al., *Virol.*, **179**:857-861, 1990; Elango et al., *J. Gen. Virol.* **69**:2893-28900, 1988.

Conditions or Diseases Amenable to Treatment Using the Method of the Invention

Various disease conditions are amenable to treatment using the methods of the invention. One skilled in the art can recognize the appropriate protein which should be produced by the invention for treating specific disease conditions. Exemplary diseases which are amenable to treatment using the subject invention, and exemplary, appropriate proteins which can be used in treating these diseases, are shown in Table 2.

TABLE 2: Exemplary Disease Conditions Amenable to Treatment Using the Invention

Enzyme Deficiency Adenosine deaminase[1] Purine nucleotide phosphorylase Galactosidase β-glucuronidase	**Endotoxic Shock/Sepsis** Lipid-binding protein (LBP)
Antioxidants for Cancer Therapy Superoxide dismutase Catalase	**Anemia** Erythropoietin
Cancer α-Interferon γ-Interferon α-IL1 Phenylalanine ammonia lyase Arginase L-asparaginase Uricase Granulocyte colony stimulating factor (G-CSF) Monoclonal antibodies Tissue necrosis factor	**Growth Factors (for use in wound healing, induction of red blood cell formation, etc.)** Epidermal growth factor G-CSF γ-Interferon Transforming growth factor Erythropoietin Thrombopoietin Insulin-like growth factor-1 Insulin Human growth hormone
Cardiovascular Disease Tissue plasminogen activator Urokinase (native or chimeric) α_1-antitrypsin Antithrombin-III Other proteases or protease inhibitors Apolipoproteins (particularly B-48) Circulating Scavenger Receptor APO A1[2]	**Diabetes** Insulin Glucagon Insulinotrophic hormone **Clotting disorders** Clotting factor VIII
Obesity and Feeding Ob gene product Cholecystokinin (CCK)	**Gastrointestinal and Pancreatic Deficiencies** Pepsin (for esophageal reflux) Trypsin Chymotrypsin Elastase Carboxypeptidase Lactase (for lactose deficiency) Sucrase: Intrinsic Factor (pernicious anemia)
Bone diseases Calcitonin PTH-like hormone	

TABLE 2 (cont.)

Organ-Specific Autoimmune diseases (target of antibody in parentheses) Myasthenia gravis (acetylcholine receptors) Graves' disease (thyroid-stimulating hormone receptor) Thyroiditis (thyroid, peroxidase) Insulin-resistant diabetes with acanthosis nigricans or with ataxia telangiectasia (Insulin receptor) Allergic rhinitis, asthma ($Beta_2$-adrenergic receptors) Juvenile insulin-dependent diabetes (insulin, GAD65) Pernicious anemia (gastric parietal cells, vitamin B_{12} binding site of intrinsic factor) Addison's disease (adrenal cells) Idiopathic hypoparathyroidism (parathyroid cells) Spontaneous infertility (sperm) Premature ovarian failure (interstitial cells, corpus luteum cells) Pemphigus (intercellular substance of skin and mucosa) Bullous pemphigoid (basement membrane zone of skin and mucosa) Primary biliary cirrhosis (mitochondria) Autoimmune hemolytic anemia (erythrocytes) Idiopathic thrombocytopenic purpura (platelet) Idiopathic neutropenia (neutrophils) Vitiligo (melanocytes) Osteosclerosis and Meniere's disease (type II collagen) Chronic active hepatitis (nuclei of hepatocytes)
Systemic Autoimmune Diseases (defect/organ affected in parentheses) Goodpasture's syndrome (basement membranes) Rheumatoid arthritis (γ-globulin, EBV-related antigens, collagen types II and III) Sjögren's syndrome (γ-globulin, SS-A (Ro), SS-B (La)) Systemic lupus erythematosus (nuclei, double-stranded DNA, single-stranded DNA, Sm ribonucleoprotein, lymphocytes, erythrocytes, neurons, γ-globulin) Scleroderm (nuclei, Scl-70, SS-A(Ro), SS-B (La), centromere) Polymyositis (nuclei, Jo-1, PL-7, histadyl-tRNA or threonyl-tRNA synthetases, PM-1, Mi-2) Rheumatic fever (myocardium heart valves, choroid plexus)
[1] For treatment of severe combined immunodeficiency
[2] Converts low-density lipoproteins to high-density lipoproteins

Transformation of Secretory Gland Cells

The DNA of interest-containing vector (*i.e.*, either a viral or non-viral vector (including naked DNA)) is introduced into the secretory gland *in vivo* via the duct system (*i.e.*, by retrograde ductal injection, which may be accomplished by perfusion (*i.e.*, continuous injection), or by a single, discontinuous injection). Retrograde ductal injection may be accomplished in the pancreas and liver by endoscopic retrograde chalangio-pancreatography (ECRP). Ductal administration provides several advantages. Because the vector is presented to the cells from "outside" the body (from the lumen), the immunological and inflammatory reactions that are commonly observed as a result of the administration of transforming formulations and their adjuvants into blood and interstitial fluid may be avoided.

Moreover, the cells of secretory glands form a monolayer that encloses the duct system. As a consequence, virtually all of the cells of the glands can be accessed by a single administration into

the duct. In this way it is possible to transfect large masses of cells in a relatively simple manner with a single procedure. The DNA of interest can thus also be administered without substantial dilution (it is only diluted by the fluid in the duct system) and without the need to develop organ specific targeting signals. In contrast, intravenous administration necessarily greatly dilutes the material and requires that it be targeted to the organ of interest in some fashion.

The amount of DNA to transform a sufficient number of secretory gland cells and provide for expression of therapeutic levels of the protein can be readily determined using an animal model (*e.g.*, a rodent (mouse or rat) or other mammalian animal model) to assess factors such as the efficiency of transformation, the levels of protein expression achieved, the susceptibility of the targeted secretory gland cells to transformation, and the amounts of DNA required to transform secretory gland cells.

The precise amount of DNA administered will vary greatly according to a number of factors including the susceptibility of the target cells to transformation, the size and weight of the subject, the levels of protein expression desired, and the condition to be treated. For example, the amount of DNA introduced into a secretory gland of a human is generally from about 1 μg to 200 mg, preferably from about 100 μg to 100 mg, more preferably from about 500 μg to 50 mg, most preferably about 10 mg. Specifically, the amount of DNA introduced into the pancreas of a human is, for example, generally from about 1 μg to 100 mg, preferably about 100 μg to 10 mg, more preferably from about 250 μg to 5 mg, still more preferably from about 500 μg to 1.5 mg, most preferably about 1 mg. The amount of DNA introduced into the salivary gland of a human is, for example, generally from about 2.5 μg to 30 mg, more preferably from about 25 μg to 3 mg, still more preferably from about 100 μg to 1 mg, most preferably about 250 μg. The amount of DNA introduced into the liver of a human is, for examples, generally from about 10 μg to 500 mg, more preferably from about 100 μg to 300 mg, still more preferably from about 150 μg to 100 mg, most preferably about 1 mg

Generally, the amounts of DNA for human therapy according to the invention can be extrapolated from the amounts of DNA effective for therapy in an animal model. For example, the amount of DNA for therapy in a human is roughly 100 times the amount of DNA effective in therapy in a rat. The amount of DNA necessary to accomplish secretory gland cell transformation will decrease with an increase in the efficiency of the transformation method used.

Concurrent Transformation of Multiple Secretory Glands

In a preferred embodiment of the invention, at least two secretory glands are transformed according to the methods of the invention. Any two secretory glands can be transformed concurrently. For example, the DNA of interest can be administered to both the pancreas and the

liver, or to both the salivary gland and the pancreas, or to both the salivary gland and the liver, or to all three. Preferably, cells of the pancreas and the liver are concurrently transformed.

Concurrent transformation of the secretory glands can be carried out several hours to several days apart or, preferably, simultaneously (*i.e.*, DNA is introduced into the two secretory glands during the same procedure. For example, where the liver and pancreas are to be concurrently transformed, the DNA formulation can be introduced simultaneously via a common duct, or separately (*e.g.*, first via the pancreatic duct with occlusion of the hepatic duct, then vice versa).

Concurrent transformation of at least two or more secretory glands can advantageously provide higher levels of expression of the polypeptide of interest in a secretory gland tissue and/or in the bloodstream and can, unexpectedly, provide for synergy between the organs (*e.g.*, to provide for higher levels of tissue expression in a secretory gland than when the secretory gland is transformed alone). For example, concurrent transformation of the pancreas and the liver results in increased levels of tissue expression in the pancreas relative to tissue levels in pancreas when it is transformed alone.

Moreover, the liver releases the polypeptide of interest in a continuous fashion that is not regulated by hormonal stimulation. The pancreas provides a relatively lower level of constitutive secretion and stores most of the polypeptide of interest and only releases large amounts after stimulation (*e.g.*, after the individual eats). Therefore, transformation of both liver and pancreas has the advantage of providing both constitutive secretion primarily form the liver, and hormonally-regulated secretion from the pancreas.

Intravenous protein therapy by transformation of salivary gland, pancreatic, and liver cells

Secretory glands transformed according to the invention facilitate high levels expression of a DNA of interest, particularly where the DNA of interest is operably linked to a strong eukaryotic promoter (*e.g.*, CMV, MMTV). The expressed protein is then secreted at high levels into the bloodstream. The protein so expressed and secreted is thus useful in treating a mammalian subject having a variety of conditions.

In a preferred embodiment, the proteins are secreted into the bloodstream at levels sufficient for intravenous protein therapy. For example, the amount of a specific protein normally released into the blood from the pancreas can be substantial, *e.g.*, a specific protein that is released into the bloodstream can be as much as 25% of the amount of duct-directed secretion of that specific protein. This amounts to as much as 1-2 mg of protein/gram of tissue being directed into the blood per hour.

Bloodstream levels of the therapeutic protein may be enhanced by several different methods. For example, bloodstream levels can be enhanced by increasing the overall level of expression of the desired protein, *e.g.*, by integration of multiple copies of the DNA of interest into the genome of the

target cells, by operably linking a strong promoter (*e.g.*, a promoter from CMV) and/or enhancer elements to the DNA of interest in the construct, or by transformation of a greater number of target cells in the subject (*e.g.*, by administration of multiple doses of the transforming material).

Secretion of the therapeutic protein into the bloodstream can also be enhanced by incorporating leader sequences, amino acid sequence motifs, or other elements that mediate intravenous-directed secretion into the sequence of the therapeutic protein. For example, the DNA of interest can be engineered to contain a secretion signal that directs secretion of the protein primarily into the bloodstream, thereby increasing the amount of the protein produced in the secretory gland that reaches in the bloodstream. Intravenous-directed secretion signals can be identified by, for example, site-directed mutagenesis of DNA encoding a bloodstream-targeted protein (*e.g.*, insulin). The mutants can be screened by expression of the mutated DNA in secretory gland cells and subsequently determining the ratio of, for example, salivary to intravenous expression.

Alternatively, intravenous-directed secretion signals can be identified by constructing recombinant, chimeric proteins composed of, for example, a putative intravenous secretion signal inserted into a saliva-directed protein. Intravenous secretion signals would then be identified by their ability to re-direct expression of the saliva-directed protein into the bloodstream. Putative intravenous secretion signals and duct system secretion signals can also be identified by comparison of DNA and amino acid sequences of proteins which are preferentially secreted into the bloodstream. Areas of homology or common motifs among the proteins could then be tested as described above.

Overall secretion from secretory glands can be augmented by hormonal stimulation. For example, where the protein is primarily secreted into the duct system and is secreted at lower levels into the bloodstream, hormonal stimulation enhances intravenous secretion as well as secretion into the duct. Thus, therapeutically effective levels of the protein the bloodstream may be achieved or enhanced by administration of an appropriate, secretory gland specific hormone. For example, secretory gland secretion can be enhanced by administration of a cholinergic agonist such as acetyl-β-methyl choline, or can be augmented or further augmented by control of diet (*i.e.*, eating stimulates pancreatic and salivary gland secretion). Thus, because eating a meal can elicit a secretory response, adjustment of meals (*e.g.*, frequency of meals and/or amounts eaten) can be used as a dosing mechanism for delivery of the desired protein, and can be accomplished without administration of additional protein-encoding DNA.

Bloodstream-directed secretion can also be regulated at either the level of transcription, translation, or secretion. Transcriptional regulation involves the timing and level of transcription directed from the DNA of interest, while translational regulation involves the production of polypeptides from transcribed RNA. Secretory regulation involves the release of polypeptides from

the cell (*e.g.*, from secretory cells in which the polypeptides to be secreted are stored within intracellular vacuoles). Methods for providing transcriptional and/or translational regulation of a DNA of interest are well known in the art (e.g., transcriptional regulation through the use of inducible promoters).

Secretory regulation can be achieved by, for example, administration of a hormone that elicits a secretory response in the desired secretory gland, or by activity that stimulates production of such hormone(s) (*e.g.*, eating to stimulate pancreatic secretion). Unlike regulation at the level of transcription or translation, which can take many hours to become effective, regulation of secretion occurs within minutes after stimulation. Moreover, endocrine secretion from the pancreas and salivary glands is stimulated by hormones and neurotransmitters that are natural components of the feeding response; thus feeding itself can act as a dosing mechanism.

The actual number of transformed secretory gland cells required to achieve therapeutic levels of the protein of interest will vary according to several factors including the protein to be expressed, the level of expression of the protein by the transformed cells, the rate of protein secretion, the partitioning of the therapeutic protein between the gastrointestinal tract and the bloodstream, and the condition to be treated. For example, the desired intravenous level of therapeutic protein can be readily calculated by determining the level of the protein present in a normal subject (for treatment of a protein deficiency), or by determining the level of protein required to effect the desired therapeutic result.

Application of the Method of the Invention to Achieve Euglycemia in a Diabetic Syndrome

In another preferred embodiment of the invention, pancreatic cells are transformed using insulin-encoding DNA to provide for expression and secretion of insulin into the bloodstream of a mammalian subject. Transformation of pancreatic cells with insulin encoding DNA not only provides for regulated expression of insulin in a mammalian subject, but also provides for maintenance of a euglycemic state (*i.e.*, normal blood glucose levels) in diabetic subjects for extended periods of time (*e.g.*, up to 6 to 7 days post transformation). Thus, not only is the exocrine pancreas secreting insulin to reduce blood sugar, but regulating its secretion so that blood levels are maintained at normal levels, e.g., are regulated. Thus, pancreatic transformation with insulin-encoding DNA can be used in the therapy of individuals having a disease or condition associated with elevated blood glucose levels (*e.g.*, diabetes (*e.g.*, type I or type II diabetes), and hyperglycemia). This aspect of the invention may be applied to regulate levels of other proteins in the bloodstream.

Assessment of Protein Therapy

The effects of expression of the protein encoded by the DNA of interest following *in vivo* transfer of the DNA of interest can be monitored in a variety of ways. Generally, a sample of blood from the subject can be assayed for the presence of the therapeutic protein. Appropriate assays for detecting a protein of interest in blood samples are well known in the art. For example, a sample of blood can be tested for the presence of the polypeptide using an antibody which specifically binds the polypeptide in an ELISA assay. This assay can be performed either qualitatively or quantitatively. The ELISA assay, as well as other immunological assays for detecting a polypeptide in a sample, are described in Antibodies: A Laboratory Manual (1988, Harlow and Lane, eds. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).

Alternatively, or in addition, the efficacy of the polypeptide therapy can be assessed by testing a sample of blood for an activity associated with the polypeptide (*e.g.*, an enzymatic activity). Furthermore, the efficacy of the therapy using the methods of the invention can be assessed by monitoring the condition of the mammalian subject for improvement. For example, where the polypeptide is erythropoietin, the subject's blood is examined for iron content or other parameters associated with anemia.

EXAMPLES

The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how to carry out the invention and is not intended to limit the scope of what the inventors regard as their invention. Efforts have been made to ensure accuracy with respect to numbers used (*e.g.*, amounts, temperatures, etc.), but some experimental error and deviation should be accounted for. Unless indicated otherwise, parts are parts by weight, molecular weight is weight average molecular weight, temperature is in degrees Centigrade, and pressure is at or near atmospheric.

Example 1: *In vivo* gene transfer of DNA encoding human growth hormone by retrograde injection of DNA into the pancreas

Four constructs for expression of human growth hormone (hGH) were prepared using techniques well known in the art (see, for example, Sambrook et al. *ibid*). The first construct, pFGH, contains the genomic hGH DNA sequence inserted in the commercially available vector pBLUESCRIPT SK+™ (Stratagene, LaJolla California) (Fig. 1). Because the hGH coding sequence is not linked to a promoter, this vector provides for no or only low-level hGH expression. Thus, the pFGH construct serves as a negative control for hGH expression in the pancreas. The second construct, pFGH.CMV, was constructed by operably inserting the promoter from the immediate

early gene of human CMV upstream of the genomic hGH sequence of the pFGH vector (Fig. 2). The third construct, pFGH.chymo, was constructed by operably inserting the rat chymotrypsin B gene promoter upstream of the genomic hGH sequence of the pFGH vector (Fig. 3). The fourth construct, pFGH.RSV, was constructed by operably inserting the promoter from the long terminal repeat (LTR) of RSV upstream of the genomic hGH sequence of the pFGH vector.

Each of the four vectors was used to transfect the pancreas of approximately 300 g adult male, Sprague-Dawley rats (pFGH + lipofectin, 4 rats; pFGH.chymo + lipofectin, 4 rats; pFGH.RSV + lipofectin, 4 rats; pFGH.CMV + lipofectin, 10 rats; pFGH.CMV without lipofectin, 7 rats; negative control (no DNA, no lipofectin), 3 rats). Pancreatic transfection was accomplished by first anesthetizing the rats and performing a laparotomy to expose the duodenum. The pancreas and the associated common bile duct were identified, and the common bile duct was cannulated either extraduodenally or through the papilla of Vater. The hepatic duct was occluded, and 100 µl of phosphate-buffered saline (PBS) containing one of the four vectors, or 100 µl of PBS alone as a negative control, were slowly injected or infused into the pancreatic duct in a retrograde direction. The vector-containing solutions were composed of 8 µg DNA per 100 µl in PBS, either with or without 6% lipofectin, a cationic lipid used to increase transfection efficiency. The solution was left in place for 5 min before secretory flow was allowed to resume and hepatic duct blockage removed. The catheter was left in place and inserted into the duodenum through a small hole to ensure adequate biliary and pancreatic flow post-operatively. The abdomen was then closed with sutures. The animals recovered fully and rapidly from the surgery without obvious side effects. This transfection method provides direct access of the vector to over 90% of the pancreatic gland cells.

At 48 hr after surgery, a blood sample was obtained to measure serum hGH levels, and the rats were sacrificed. At autopsy, the pancreas of both control and test rats appeared normal, and exhibited no gross or microscopic pathology.

The pancreas was dissected free from the mesenteric surface and was homogenized in cold 0.2 M (pH 8.0) sodium phosphate buffer (1:10 w/v) containing protease inhibitors aprotinin, leupeptin, pepstatin, and PEFABLOC SC™. Homogenization was completed by shearing after 10 passes with a motorized pestle at approximately 4000 rpm in a glass homogenizer. The homogenate was then centrifuged at 1000 g for 15 min. The supernatant was collected and stored at -80°C until analysis. The levels of hGH in the serum and pancreatic protein samples were measured using the hGH radioimmune assay (Nichols Institute). Each assay was performed in duplicate and compared to a set of control samples.

Rats injected with the pFGH.CMV vector expressed higher levels of hGH in the pancreatic tissue (Fig. 4), compared to background levels of pancreatic hGH expression in rats injected with either no DNA (PBS alone) or the pFGH vector (hGH DNA with no promoter). The addition of

lipofectin modestly increased hGH expression in rats injected with the pFGH.CMV construct. In addition, rats transfected with the pFGH.CMV vector secreted hGH in the serum at levels increased relative to background levels and to hGH secretion levels in rats injected with either control samples (no DNA or pFGH) or with samples containing hGH DNA linked to either the chymotrypsin B or RSV promoters (Fig. 5). In Fig. 6, all data from the above experiments (including all controls and vectors) are analyzed by plotting the hGH serum levels against the hGH tissue levels. This graph shows that higher tissue levels result in higher levels of secretion into the blood. Thus, retrograde pancreatic injection of the pFGH.CMV vector successfully transfected pancreatic cells to provide both hGH pancreatic tissue expression and hGH secretion into the bloodstream.

Example 2: *In Vivo* Transformation of Pancreatic Cells by Retrograde Ductal Injection of hGH-Encoding DNA and Regulation of hGH Secretion

Eight rats were anesthetized and control blood samples (no DNA) were collected from the femoral vein of each animal. Pancreatic transfection was accomplished by exposing the duodenum by laparotomy and identifying the pancreas and the associated common bile duct. The common bile duct was cannulated either extraduodenally or through the papilla of Vater, and the hepatic duct was occluded. A 1:50 dilution of replication-defective human adenovirus (Ad5-*dl* 342) supernatant in 100 µl of phosphate-buffered saline (PBS) containing 8 µg of the hGH-encoding plasmid pFGH.CMV (Fig. 2) was slowly infused into the pancreatic duct in a retrograde direction. The solution was left in place for approximately 5 min before secretory flow was allowed to resume and the hepatic duct blockage removed. The catheter was left in place and inserted into the duodenum through a small hole to ensure adequate biliary and pancreatic flow post-operatively. The abdomen was then closed with sutures. The animals recovered fully and rapidly from the surgery without obvious side effects.

At 48 hr after surgery, a blood sample was obtained to measure serum hGH levels (unstimulated serum levels). The cholinergic agonist McH was injected subcutaneously into each rat at 0.8 mg/kg body weight. Blood samples were collected from the inferior vena cava of each animal at 15 min intervals following McH injection. Serum was separated from the blood of all samples after clotting, and kept at -20°C prior to assay.

As shown in Fig. 9 (one representative animal) plasma levels of hGH increased markedly following McH injection, demonstrating that secretion of hGH expressed by transformed pancreatic cells is regulated by agonist stimulation. Moreover, bloodstream-directed secretion of hGH from the transformed pancreatic cells occurred at relevant, physiological levels useful in therapeutic administration (*i.e.*, at the ng/ml level).

Example 3: *In Vivo* Transformation of Salivary Glands by Retrograde Ductal Injection of DNA Encoding Human Growth Hormone

Twelve adult rats weighing approximately 300 g each were anesthetized with an intraperitoneal injection of sodium pentobarbital. A total volume of 50 µl containing 4 µg of the pFGH.CMV plasmid, which contains cDNA encoding human growth hormone (hGH) (Fig. 2), was introduced into each submandibular gland of 8 rats by retrograde ductal injection via the ducts leading from the oral mucosa to the salivary gland. Briefly, both the left and right Wharton's duct were cannulated intraorally with polyethylene (PE) 10 tubing, and the DNA injected into the duct system of each gland in a retrograde fashion (4 µg/50 µl of PBS). The material was kept in place for two minutes before normal flow was reestablished.

For three of these animals the DNA was mixed in a 6% solution of the cationic lipid Lipofectin (labeled "liposomes") from Life Technologies (Gaithersburg, MD). For four of these animals, the DNA was mixed with a 1:50 dilution of replication-defective human adenovirus (Ad5-*dl* 342) supernatant. Control rats (4 rats) received 50 µl 0.9% saline (control) without plasmid. No significant leakage of material or bleeding occurred. After 3 hours, the animals were awake, drinking water, and appearing normal.

Approximately 48 hours after cDNA injection, the animals were sacrificed. The right and left submandibular glands were removed and were homogenized in cold 0.2 M (pH 8.0) sodium phosphate buffer (1:10 w/v) containing the protease inhibitors aprotinin, leupeptin, pepstatin, and PEFABLOC SC™. Homogenization was completed by shearing after 10 passes with a motorized pestle at approximately 4000 rpm in a glass homogenizer. The homogenates were centrifuged at 1000 g for 15 min, and the supernatant collected and stored at -80°C until analysis. The levels of hGH in the protein samples were measured using the hGH radioimmune assay (Nichols Institute). Each assay was performed in duplicate and compared to a set of control samples.

Each of the submandibular glands of the rats injected with the pFGH.CMV vector expressed hGH in the salivary gland tissue; hGH expression was undetectable in the control rats' salivary glands (Fig. 7).

Example 4: *In Vivo* Transformation of Salivary Glands by Retrograde Ductal Injection of hGH-Encoding DNA and Regulation of hGH Secretion

Three adult rats weighing approximately 300 g each were anesthetized with an intraperitoneal injection of sodium pentobarbital. A control blood sample (prior to DNA) was drawn from the femoral vein of each animal. A total of 4 µg of the hGH-encoding plasmid pFGH.CMV (Fig. 2) in 50 µl, was introduced into each submandibular gland of each rat by retrograde ductal

injection via the ducts leading from the oral mucosa to the salivary gland as described above in Example 2. No significant leakage of material or bleeding occurred. After 3 hours, the animals were awake, drinking water, and appearing normal.

Forty-eight hours after cDNA injection, the animals were again anesthetized and a control blood sample was drawn from the femoral vein of each animal (unstimulated serum level). The cholinergic agonist acetyl-β-methyl choline (McH) was injected subcutaneously at 0.8 mg/kg body weight into each animal. Blood samples were collected from the femoral vein of each animal at 10 min, 20 min, 40 min, and 50 min after McH injection. Serum was separated from the blood of all samples after clotting, and kept at -20°C prior to assay.

As shown in Fig. 8 (one representative animal), secretion of hGH into the bloodstream was dramatically increased in response to administration of McH, peaking at 40 min. Thus, these data demonstrate that introduction of hGH-encoding DNA into the salivary gland results in bloodstream-directed secretion of hGH and regulation by cholinergic stimulation. Moreover, regulation is at the level of secretion, not transcription, since transcriptional regulation would not result in increased hGH bloodstream levels in such a short period.

Example 5: Treatment of Diabetes Mellitus Over a Three Day Period by *In Vivo* Transformation of Pancreatic Cells by Retrograde Ductal Injection with Insulin-Encoding DNA

Streptozotocin, which induces diabetes mellitus in rats, was administered to 8 male Sprague-Dawley rats (260-280 g) after overnight fasting by intraperitoneal injection in 1 mM citrate buffer (pH 4.5) (Sigma) at 65 mg/kg of body weight. One hour later, animals were anesthetized with Nembutal and the body cavity opened to expose the gastrointestinal tract. Each animal was given the appropriate DNA construct directly by retrograde injection in the pancreatic duct in a 100 µl injection volume containing 8 µg DNA plus adenovirus (Ad5-*dl* 342)(3×10^{10} viral particles) as described above. Test animals (4 rats) received the human insulin-encoding construct pBat16.hInsG1.M2. The pBAT16.hInsG1.M2 construct (Fig. 10) encodes an insulin gene containing a site-directed mutation of the second protease site to create a furin recognition site; this construct provides for enhanced expression of processed insulin in non-neuroendocrine cells. In addition, the human β-globin first intron replaces the first insulin gene intron which is inefficiently spliced. Control animals (4 rats) received the control construct CMV-GFP, which contains a green fluorescent protein (GFP)-encoding sequence operably linked to a CMV promoter. The animals recovered fully and rapidly from the surgery without obvious side effects. Body weight and blood glucose were monitored daily for three days post-injection. Blood glucose was measured by the glucose oxidase method (Lifescan, Milpitas, CA).

As shown in Fig. 11, treatment of the streptozotocin-induced diabetic rats with the insulin-encoding construct resulted in maintenance of almost complete euglycemia for 3 days. In contrast, control animals that received the GFP-encoding construct remained hyperglycemic throughout the test period. The data show that introduction of insulin-encoding DNA into the pancreas results in pancreatic cell transformation, as well as secretion of insulin by the transformed pancreatic cells at levels sufficient to overcome diabetes in an animal model. Moreover, these results show that the method of the invention provides regulated and relatively normal blood glucose levels. Surprising, the exocrine pancreas regulates the release of insulin such that blood sugar levels are maintained at regulated levels (normally the endocrine pancreas is responsible for regulation of bloodstream-directed secretion).

Example 6: Treatment of Diabetes Mellitus Over a Six Day Period by *In Vivo* Transformation of Pancreatic Cells by Retrograde Ductal Injection with Insulin-Encoding DNA

Streptozotocin was administered to 14 rats at 70 mg/kg body weight by intraperitoneal injection to induce diabetes mellitus. The animals were then anesthetized by intraperitoneal injection of sodium pentobarbital. Two rats did not receive streptozotocin and served as one negative control. Insulin-encoding DNA in the pBAT16.hInsG1.M2 construct (Fig. 10) was administered to 8 of the streptozotocin-injected rats by retrograde ductal injection as described above. Six streptozotocin-treated rats received either 100 µl of saline without DNA (2 animals) or a control DNA without the human insulin gene (4 animals) by pancreatic retrograde ductal injection as additional negative controls. The animals recovered fully and rapidly from the surgery without obvious side effects. Blood samples were collected from the femoral vein of each animal at 24 hr intervals for 6 days. Human insulin was measured using a double antibody radioimmunoassay (Linco Laboratories, Saint Louis, MO).

As shown in (Fig. 12), blood glucose levels were significantly decreased in the diabetic rats that received the insulin-encoding DNA (+Strep, + DNA) relative to diabetic the rats that received no DNA (+Strep, No DNA). Furthermore, these decreased blood glucose levels were observed throughout the entire 6 day course of the experiment. Thus, these data show that introduction of insulin-encoding DNA into the pancreas results in persistent expression of insulin, and that the insulin expressed by the transformed pancreatic cells is secreted into the bloodstream and can function in regulation of blood glucose at levels sufficient to overcome diabetes in an animal model. As shown in Fig. 12, elevated insulin levels for such an extended period additionally demonstrate prolonged expression from the DNA introduced into the pancreatic cells.

Example 7: *In Vivo* Transformation of Pancreatic Cells by Retrograde Ductal Injection of Green Fluorescent Protein-Encoding DNA and Expression in Pancreatic Cells

To identify the pancreatic cells that expressed the recombinant protein, DNA encoding green fluorescent protein (GFP) was used to transform pancreatic cells according to the methods of the invention. EGFP cDNA from plasmid pEGFP.C2 (Clontech) was inserted into pFOX. The EGFP sequence was modified to contain an SV40 nuclear localization signal, in-frame at the 3' end. This addition allowed for partial nuclear localization and facilitated immunohistochemical detection. The CMV immediate early promoter was positioned upstream of the first intron of human β-globin to create the expression vector pFOX.EGFP.N2.CMV.

After fasting overnight, Male Sprague-Dawley rats (260-280 g) were anesthetized and the body cavity opened to expose the gastrointestinal tract. The green fluorescent protein (GFP)-encoding construct pFOX.EGFP.N2.CMV was administered to each animal by retrograde injection in the pancreatic duct in a 100 µl injection volume containing 8 µg DNA premixed with adenovirus (3×10^{10} viral particles) as described above. The animals recovered fully and rapidly from the surgery without obvious side effects.

Seventy-two hours post-treatment, the animals were sacrificed, and pancreases were removed and weighed (wet weight). Samples of each pancreas were fixed in 5% buffered formalin for 24-48 hours at room temperature. Fixed tissues were dehydrated and imbedded in paraffin, and 5 µm sections were processed for immunohistochemistry using standard techniques. Endogenous peroxidase was quenched in 0.7% H_2O_2/MeOH, and antigen retrieval was performed using Citra solution (Biogenex, San Ramon, CA) according to the manufacturers' instructions. Sections were preincubated for 30 minutes in 5% goat serum/phosphate-buffered saline (PBS), and then incubated overnight at 4°C with primary antisera diluted in 5% goat serum/PBS.

The primary antisera were selected from either anti-GFP antisera (1:1500; Clontech, Palo Alto, CA), anti-insulin antisera (1:500; Dako, Carpenteria, CA), or non-specific rabbit sera (1:1500). The following day all sections were incubated with biotinylated goat anti-rabbit antiserum (5 µg/ml; Vector, Burlingame, CA) for 30 minutes at room temperature, and then incubated with streptavidin-aminohexanol-biotin horseradish peroxidase (HRP) complex (Vectastain-Elite, Vector). Protein was visualized by reaction with the peroxidase substrate 3,3-diamino-benzidine tetrahydrochloride (DAB; Sigma). The color reaction was followed by a brief counter stain in 1% methyl green (Sigma) prior to mounting. Negative controls included staining of sections from pancreas not injected with CMV-GFP, and omission of primary antiserum.

Staining for GFP was observed in the pancreas of animals treated with GFP DNA, but not in control animals. GFP expression was restricted to exocrine cells; there was no staining in either ductal or islet cells. Moreover, expression was observed in 0.1-1.0% of exocrine cells. Endogenous

insulin was detected in adjacent sections; but GFP expression did not co-localize with insulin expression, suggesting that the pancreatic cells primarily transformed are exocrine, not endocrine cells. Under the conditions studied there was no histological indication of inflammatory infiltration as a consequence of ductal injection of the vector.

These data show that introduction of the DNA construct results in successful transformation of pancreatic cells, despite the introduction of the construct against the flow of pancreatic juices and the high concentrations of DNase in the pancreatic juice. Moreover, these data, combined with the data above showing that transformation of the pancreas results in bloodstream-directed secretion of the encoded protein, and suggest that transformation of exocrine pancreatic cells results in bloodstream-directed secretion of the protein encoded by the introduced construct. Furthermore, because insulin staining and GFP staining did not co-localize, introduction of the GFP-encoding construct resulted in transformation of exocrine tissue, which is normally associated with protein secretion into the gastrointestinal tract, rather than endocrine tissue, which is normally associated with bloodstream-directed secretion. Despite this, bloodstream-directed secretion was still obtained at physiologically relevant levels sufficient to treat diabetes mellitus in an animal model as evidenced in the examples above.

Example 8: *In Vivo* Transformation of Liver Cells by Retrograde Ductal Injection of hGH-Encoding DNA and Bloodstream-Directed hGH Secretion

Four rats were anesthetized and control blood samples (no DNA) were collected from the femoral vein of each animal. Transfection of liver cells was accomplished by exposing the duodenum by laparotomy and identifying the liver and the associated common bile duct. The common bile duct was cannulated either extraduodenally or through the papilla of Vater. The tubing was advanced to the bifurcation of the hepatic duct in order to prevent injected material from entering the distally located pancreatic drainage. A 1:50 dilution of replication-defective human adenovirus supernatant in 100 µl of phosphate-buffered saline (PBS) containing 8 µg of the hGH-encoding plasmid pFGH.CMV (Fig. 2) or 100 µl of PBS alone (no DNA) were slowly infused into the hepatic duct in a retrograde direction. The solution was left in place for approximately 2 min to 5 min before secretory flow was allowed to resume and the pancreatic duct blockage removed. The catheter was left in place and inserted into the duodenum through a small hole to ensure adequate biliary and pancreatic flow post-operatively. The abdomen was then closed with sutures. The animals recovered fully and rapidly from the surgery without obvious side effects.

Plasma hGH levels were measured 2 days after treatment; the results are shown in Fig. 13. Each data point in Fig. 13 represents the mean ± standard error of the mean (SEM) for three animals. These data demonstrate that liver cells were transformed with the hGH-encoding DNA.

Furthermore, hGH was secreted by the transformed liver cells into the bloodstream at physiologically relevant levels.

Example 9: *In Vivo* Transformation of Pancreatic Cells with hGH-Encoding DNA and Expression in Rat Exocrine Pancreas and Plasma

Following overnight fasting and anesthesia with pentobarbital, the abdominal cavity of the rats was opened and the pancreatic duct cannulated external to the duodenum with PE 10 tubing as described above. Eight to twenty-five micrograms of each of pFGH (promoter less construct), pFGH.chymo (construct with the chymotrypsin promoter), pFGH.RSV (construct with the RSV promoter), and pFGH.CMV (construct with the CMV promoter) was injected in a total volume of 100 µl of PBS into the pancreas via the pancreatic duct as described above. Immediately prior to injection construct samples were optionally premixed with either Lipofectin (6-12% vol:vol) or adenovirus (3×10^{10} viral particles). The material was kept in the duct for 5 min prior to establishing normal flow. The abdomen was the closed and the animals allowed to recover.

Forty-eight hours later the pancreas was harvested, plasma obtained, and human growth hormone measured. The animals were anesthetized, blood samples taken (either from the femoral vein or inferior vena cava), and the transfected tissue removed. The tissue was homogenized in PBS containing 5 mM $Na2HPO_4$ (pH 7.8) at a tissue to fluid ratio of 1:10 using a motorized mortar and pestle. Large particulate material in the homogenate was removed by sedimentation at 10,000 x g for 30 minutes, and the supernatant assayed for the protein of interest. The results are shown in Figs. 14-17. All data shown are the mean ± the SEM.

The effects of the various promoters upon tissue expression and secretion of hGH into the bloodstream are shown in Figs. 14 and 15, respectively. In these experiments, the constructs were mixed with lipofection prior to administration. Of the promoters tested, the CMV promoter was by far the most effective, and produced high levels of hGH in tissue (in the range of 150 ng/g tissue wet weight) when compared to either promoter less controls, or plasmids containing RSV and chymotrypsin promoters (Fig. 14). The cationic lipid adjuvant Lipofectin increased expression by about 50%, and pre-mixing the plasmid with adenovirus enhanced tissue expression five fold (Fig. 15). Expression of hGH at 24, 48 or 72 hours after injection was similar under all conditions studied.

As shown in Figs. 16 and 17, hGH was secreted into plasma. Plasmids containing the CMV promoter increased circulating levels of hGH five times above background (Fig. 16). With plasmid alone, plasma hGH concentrations in the range of 60 to 80 pg/ml were routinely observed. Premixing the plasmids with adjuvants also increased circulating hGH levels (Fig. 17). Lipofectin

increased plasma levels by an additional 50%, and adenovirus by 75%, when compared to plasmid alone.

Example 10: Comparison of hGH Secretion by Rat Liver, Pancreas, and Combined Liver and Pancreas Transformed with hGH-Encoding DNA

Eight micrograms of the pFGH.CMV construct premixed with adenovirus as described above in Example 9, was injected into the ducts of either the liver, the pancreas, or both organs of the same animal. Where only the liver or the pancreas was transformed (liver alone or pancreas alone), the DNA was introduced according to the methods described above. Where both the liver and pancreas were transformed, the DNA-containing formulation was introduced into the hepatic duct first, and then the tubing partially withdrawn to provide access to the pancreatic duct system. A temporary ligature was then placed around the hepatic duct to prevent the second infusion from entering the parenchyma of the liver. Thus, animals in which both the pancreas and liver were transformed received two doses of the DNA-containing formulation. Plasma hGH levels were measured two days later.

In animals having transformed liver (liver alone) or pancreas (pancreas alone), hGH was expressed in liver or pancreatic tissue, respectively, and hGH detected in plasma under both circumstances. Tissue levels in liver when transformed alone were far lower than in the pancreas when transformed alone (less than 1 ng/g, as compared to about 500 ng/g), but hGH concentration in plasma of animals in which only the liver was transformed was nonetheless comparable to hGH plasma levels in animals having only the pancreas transformed (in the range of 0.15 ng/ml; Fig. 18). These results are consistent with the observation that, in contrast to the exocrine cells of the pancreas and salivary glands, hepatocytes secrete most of what they produce soon after synthesis.

When pancreas and liver were both transfected, plasma levels were higher than seen when the glands were treated individually (nearly 0.3 ng/ml) -- a value approximately equal to the sum of that observed for the two organs separately. Surprisingly, transformation of both liver and pancreas resulted in tissue levels in the pancreas being significantly increased relative to tissue levels in the pancreas when the pancreas was transformed alone (Fig. 19)

Example 11: Human growth hormone (hGH) expression in rat salivary gland.

Four micrograms of the pFGH.CMV construct, premixed with either Lipofectin or adenovirus, was injected into each submandibular gland via retrograde ductal injection (via Wharton's duct) as described above. Two days later, each gland was harvested and hGH content was measured

As shown in Fig. 20, tissue levels of hGH averaged about 50 ng/g tissue wet weight. Plasma hGH levels were in the 20-40 pg/ml range. As in the pancreas, the addition of adenovirus increased tissue hGH levels, in this case to 100 ng/g (Fig. 20).

Example 12: Stimulation of Human growth hormone (hGH) secretion.

Even when exocrine secretory cells store large amounts of protein, such as after a period of fasting, they secrete these proteins at a low rate under unstimulated conditions (*i.e.* basal or constitutive secretion). Greater rates are achieved when exogenous stimulants (*e.g.*, hormonal stimulants and/or stimulation associated with eating) are applied. To determine whether secretion of the engineered protein would be enhanced during feeding, pancreatic secretion was stimulated with a secretory stimulant. For these experiments we used animals in which both pancreas and liver were transfected. Eight micrograms of the pFGH.CMV construct were injected into ducts of both the pancreas and liver of four rats as described above. A blood sample was taken prior to injection as a control. Two days after transfection, a second control blood sample was taken and the rats were treated with the cholinergic agonist, acetyl-β-methylcholine (MeH) (0.8 mg/kg body weight).

As shown in Fig. 21, hGH secretion was increased three fold within 30 minutes of stimulation, with plasma levels approaching 1.0 ng/ml. Similar enhancement of hGH secretion was observed when either the pancreas was studied alone, or when the salivary glands were studied alone. These data show that hGH secretion is enhanced by stimulation with a cholinergic agonist. Thus secretion of hGH is regulated in a manner similar to secretion of endogenous proteins.

Although the concentration of hGH in plasma was correlated to the level of hormone in the pancreas ($r=0.55$, $p<0.01$, $n=41$), at high tissue levels, plasma concentration was not linearly proportional to tissue content. For example, addition of adenovirus to the hGH vector produced a five fold increase in tissue levels relative to the plasmid alone (Fig. 20), but only about a two fold increase in plasma concentration (see, *e.g.*, Figs. 16 and 17, Example 9 above). This lack of proportionality indicates that it is not the concentration of product in the cells alone that determines the rate of secretion into blood, but that at high tissue levels, secretion is limited by other factors. This result is similar to what is observed for endogenous protein secretion and suggests that secretion of the engineered protein is regulated in much the same manner.

Example 13: Human Insulin Expression and Secretion in Diabetic Rat Pancreas

In an attempt to treat a disease state, diabetes mellitus, we expressed human insulin in the exocrine pancreas. Fasted experimental and control animals received intra-peritoneal streptozotocin (Sigma; 65 mg/kg body weight, in 1mM citrate buffer, pH4.5) on day zero one hour prior to administration of the insulin-encoding construct. The experimental animals subsequently received 8

μg of the insulin plasmid (pBAT16.hInsG1.M2) premixed with adenovirus and injected into the pancreatic duct, also on day zero. The pBAT16.hInsG1.M2 construct contains the human insulin cDNA linked to a CMV immediate early promoter, which is positioned upstream of the first intron of human β-globin. The human insulin cDNA was mutated to convert the second protease site, between peptides C and A, to a furin recognition site. This allows for correct proteolytic processing of mature insulin in non-endocrine cells.

Plasma insulin and glucose levels were determined for up to six days. Plasma glucose levels in diabetic rats (n=3), and diabetic rats treated with the pBAT16.hInsG1.M2 plasmid (n=3), measured over a three day period, are shown in Fig. 22. Plasma insulin levels in diabetic rats (n=3), and diabetic rats treated with the pBAT16.hInsG1.M2 plasmid (n = 3), measured over a three day period, are shown in Fig. 23. Plasma glucose levels in individual diabetic (n=3) and pBAT16.hInsG1.M2 plasmid- treated diabetic rats (n=3), measured over a six day period, are shown in Fig. 24.

As a consequence of streptozotocin administration, blood glucose levels rose from the normal level of 100 mg/dl to 300-400 mg/dl within 24 hours and remained elevated for the duration of the study (Fig. 22). Treatment with the human insulin plasmid reduced blood glucose levels in diabetic rats to the normal range (Figs. 22 and 24), and concentrations of insulin remained near pre-treatment values (Fig. 23). Blood glucose levels were euglycemic for the duration of the study (6 days; Fig. 24). Animals transfected with a control plasmid remained diabetic (data not shown). These data show that regulation of insulin secretion in response to feeding was effective.

Example 14: *In vivo* gene transfer of DNA encoding human growth hormone by retrograde injection of DNA into the salivary gland

A DNA expression construct encoding human growth hormone (hGH) is prepared by operably linking a CMV promoter to hGH-encoding DNA. The expression cassette is then inserted into a construct such as the bacterial plasmid pBR322. *Escherichia coli* is then transformed with the plasmid using conventional transformation procedures. *E. coli* containing the plasmid are selected by virtue of the tetracycline or ampicillin resistance encoded by pBR322, and the transformed bacterial cells propagated in culture. Plasmid DNA is then isolated from the transformed bacterial cell culture and the DNA purified by cesium gradient.

Approximately 250 μg of the purified plasmid DNA containing hGH DNA is injected into the salivary gland of a human patient by retrograde ductal injection via a salivary gland duct. Expression and intravenous secretion of the protein is assessed using the method described above.

Example 15: *In vivo* gene transfer of DNA encoding human growth hormone by retrograde ductal injection of naked DNA into the pancreas

A construct containing hGH-encoding DNA (Marshall et al., *Biotechnology* 24:293-298, 1992) operably linked to the CMV promoter is resuspended in 0.9% saline and a volume of the DNA solution is administered to a human patient. Approximately 1 mg of DNA is delivered to the pancreas of the patient by cannulation of the pancreatic duct by duodenal intubation using endoscopic retrograde cholangio-pancreatography. Expression and secretion of human growth hormone into the bloodstream is assessed by detection of the protein in the patient's blood.

Example 16: *In vivo* gene transfer of DNA encoding human insulin by cannulation of naked DNA into the liver

A construct containing human insulin-encoding DNA operably linked to the CMV promoter is resuspended in 0.9% saline and a volume of the DNA solution is administered to a human patient. Approximately 1 mg of DNA is delivered to the patient's liver by cannulation of the hepatic duct. Expression and secretion of human growth hormone into the bloodstream is assessed by detection of the protein in the patient's blood.

Example 17: *In vivo* gene transfer of human insulin-encoding DNA to both the pancreas and the liver of a patient

A construct containing human insulin-encoding DNA operably linked to the CMV promoter is resuspended in 0.9% saline. A volume of the DNA solution is administered to a human patient so as to transform both pancreatic and hepatic cells (*e.g.*, by introducing the DNA solution into the common bile duct before it splits into the hepatic and pancreatic ducts). Approximately 2 mg of DNA is delivered to the patient's liver by cannulation of the hepatic duct; in addition, approximately 1 mg of DNA is delivered to the pancreas via retrograde injection via the pancreatic duct. Expression and secretion of human growth hormone into the bloodstream is assessed by detection of the protein in the patient's blood.

Following procedures similar to those described above, other therapeutic proteins can be expressed from DNA inserted in the genome of a secretory gland cell by gene transfer according to the invention.

The invention now being fully described, it will be apparent to one of ordinary skill in the art that many changes and modifications can be made thereto without departing from the spirit or scope of the appended claims.

What is claimed is:

1. A method of delivering a polypeptide to a mammalian subject, the method comprising:

introducing *in vivo* into at least two secretory glands of a mammal a construct comprising a DNA of interest that encodes a desired polypeptide and a eukaryotic promoting sequence operably linked to the DNA of interest, said introducing being by retrograde ductal injection;

wherein the introduced construct genetically transforms cells of each secretory gland and wherein expression of the DNA by the genetically transformed cells of the secretory glands results in bloodstream-directed secretion of the protein encoded by the DNA of interest.

2. The method of claim 1, wherein the secretory glands are selected from the group consisting of pancreas, liver, and salivary gland.

3. The method of claim 1, wherein the secretory glands are liver and pancreas.

4. The method of claim 1, wherein the DNA of interest encodes a polypeptide selected from the group consisting of insulin, growth hormone, interferon-alpha 2b, Interferon-alpha 2a, interferon-alpha N1, filgastim, insulinotropin, imiglucerase, clotting factor VIII, interferon-beta1b, erythropoietin, sargramostim, interleukin-2, interferon-gamma, anti-CD3 antibody, GPIIb/IIIa monoclonal antibody, adenosine deaminase, interleukin-8, insulin-like growth factor-1, platelet-derived growth factor, epidermal growth factor, hemoglobin.

5. The method of claim 1, wherein the DNA of interest encodes human insulin.

6. A method of controlling blood glucose levels in a mammalian subject having a disease or condition associated with hyperglycemia, the method comprising:

introducing into a mammalian patient's pancreas a construct comprising insulin-encoding DNA and a eukaryotic promoting sequence operably linked to the insulin-encoding, said introducing being by intraductal injection, thereby accomplishing genetic transformation of a pancreatic cell; and

allowing the genetically transformed cell to express the protein and secrete the protein into the bloodstream of the patient in a therapeutically effective amount thereby treating the patient;

wherein blood glucose levels in the mammalian subject are decreased to achieve a euglycemic state.

7. The method of claim 6, wherein the euglycemic state is maintained for at least three days.

8. The method of claim 6, wherein the euglycemic state is maintained for at least 6 days.

9. The method of claim 6, the method further comprising enhancing insulin expression by administration of a cholinergic agonist to the mammalian subject.

10. The method of claim 6, the method further comprising enhancing insulin expression by ingestion of a meal.

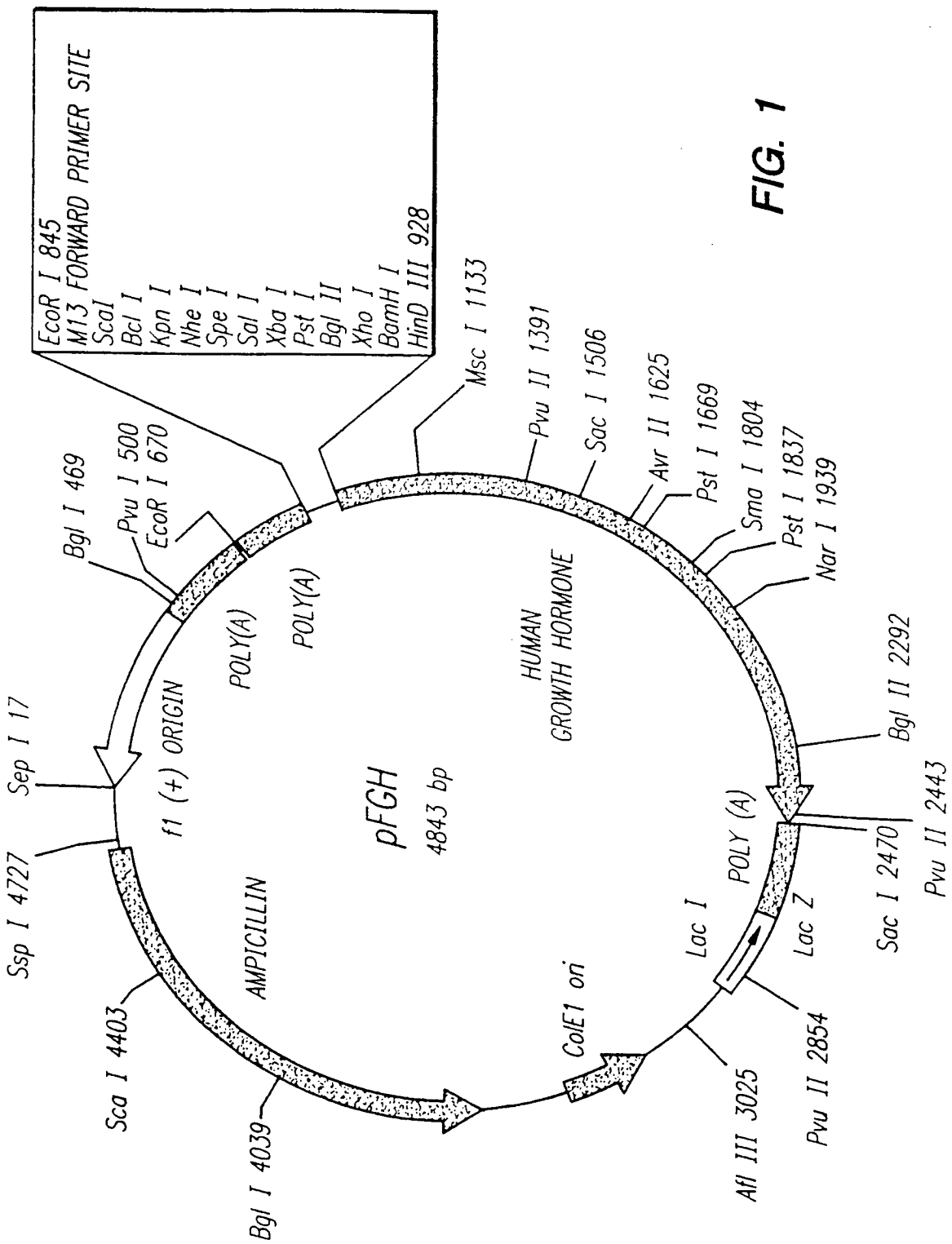


FIG. 1

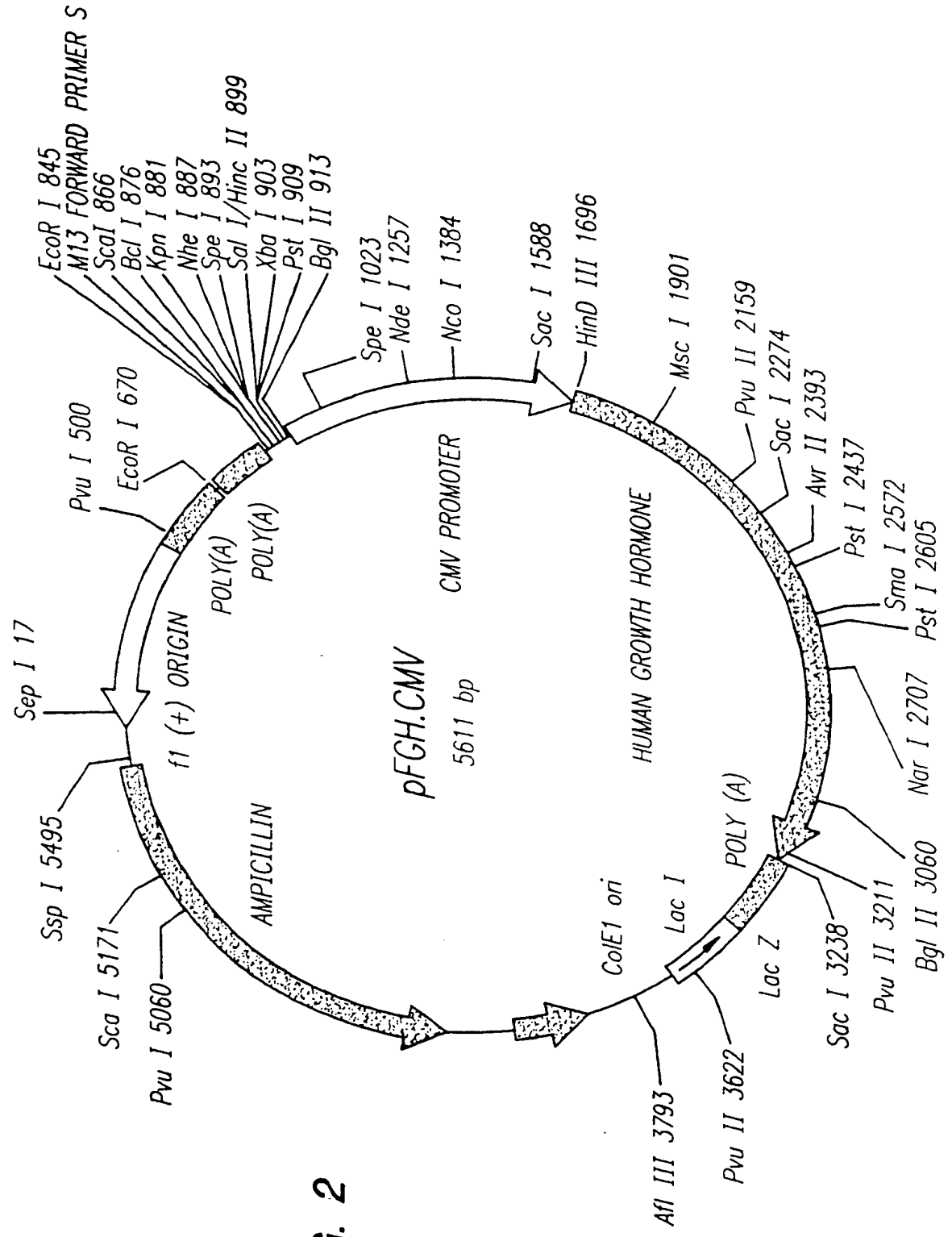


FIG. 2

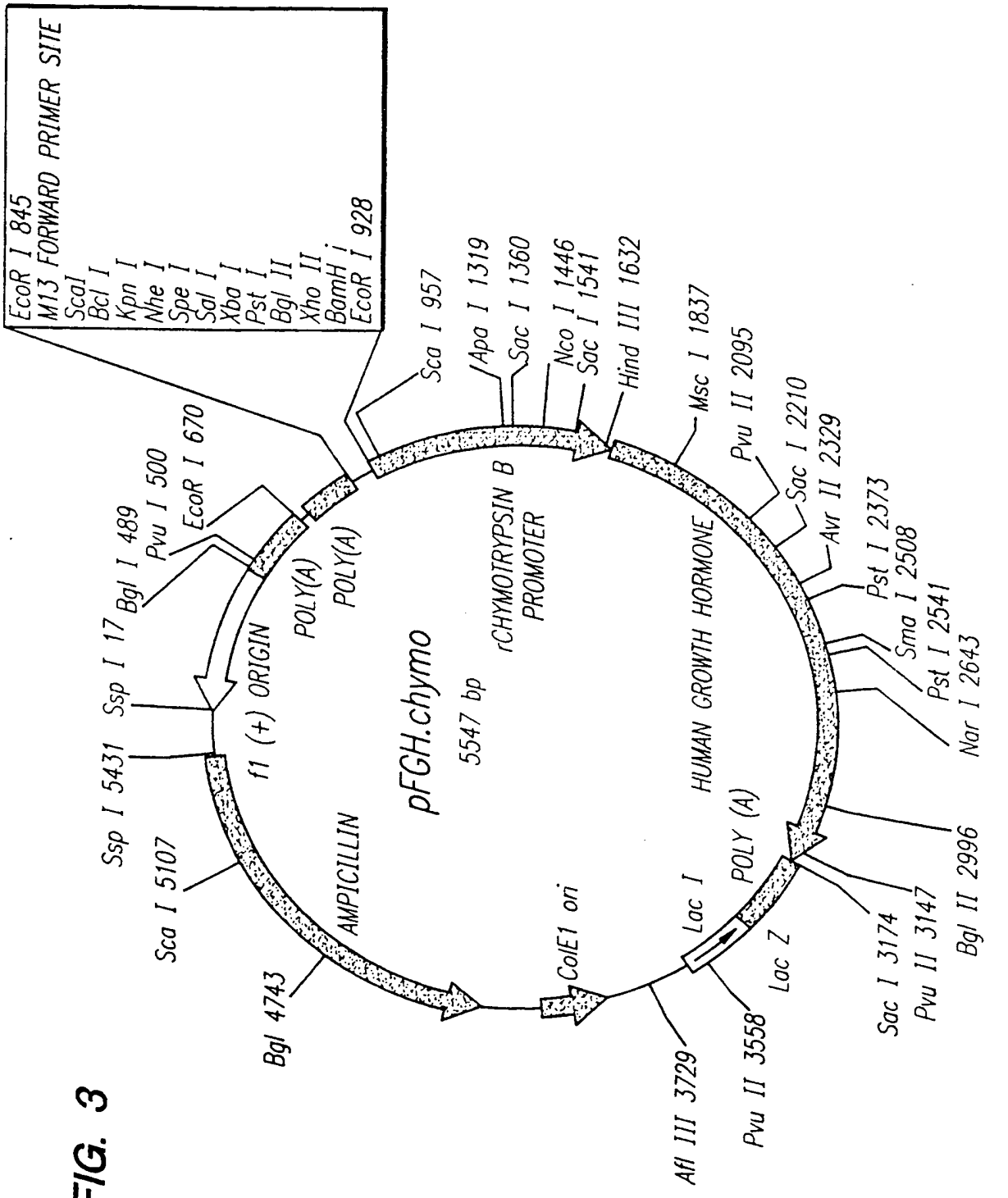


FIG. 3

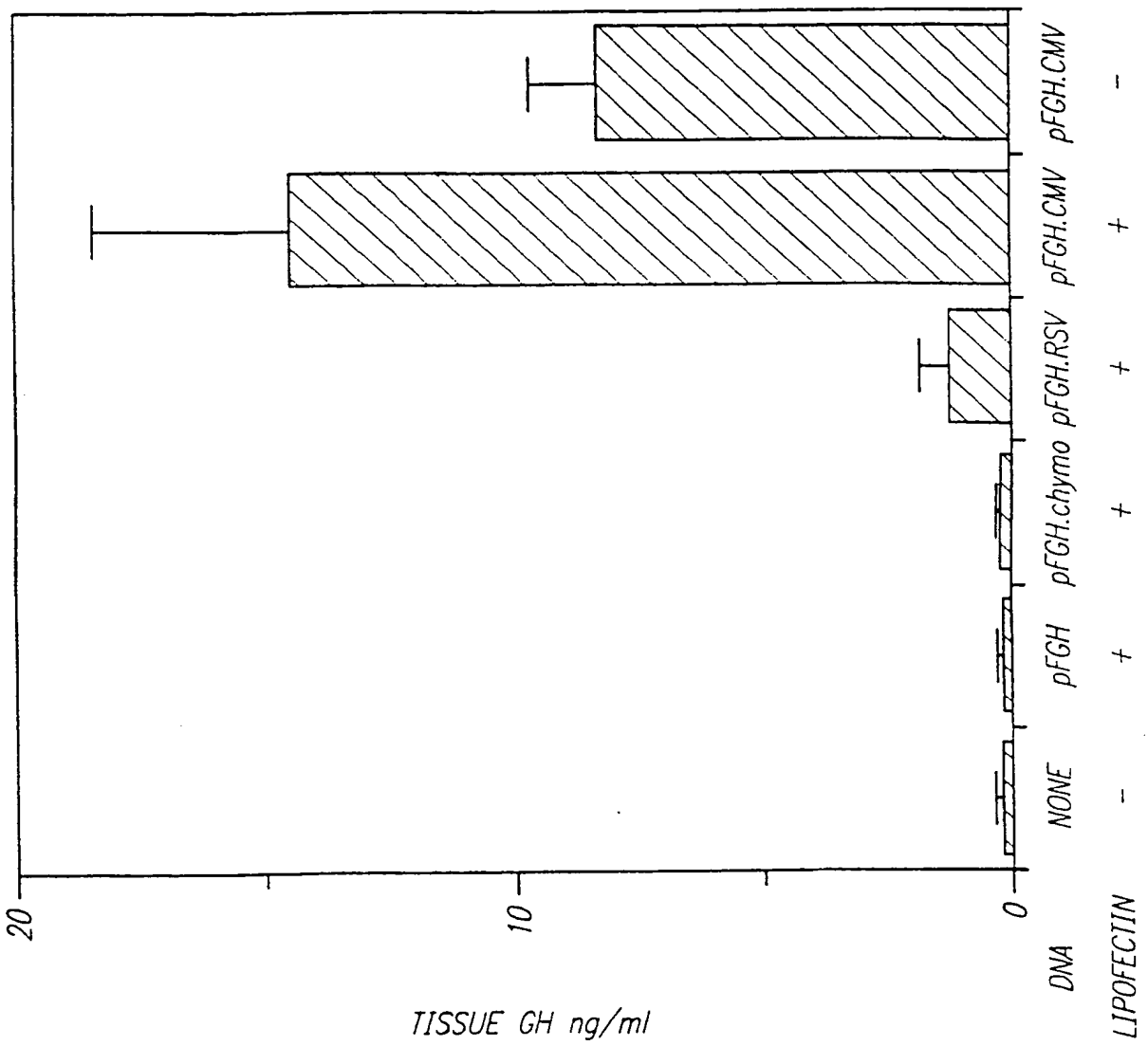


FIG. 4

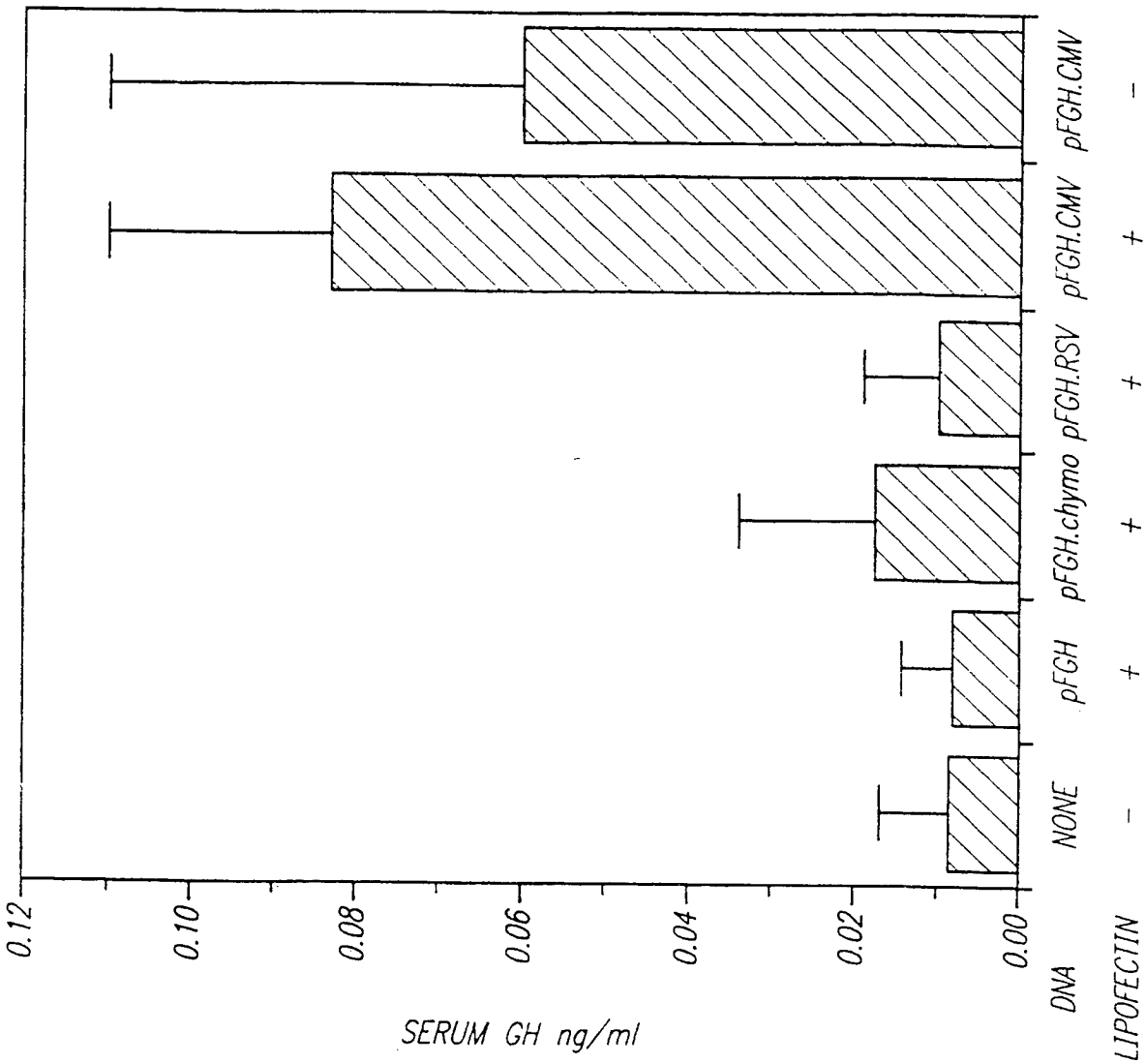


FIG. 5

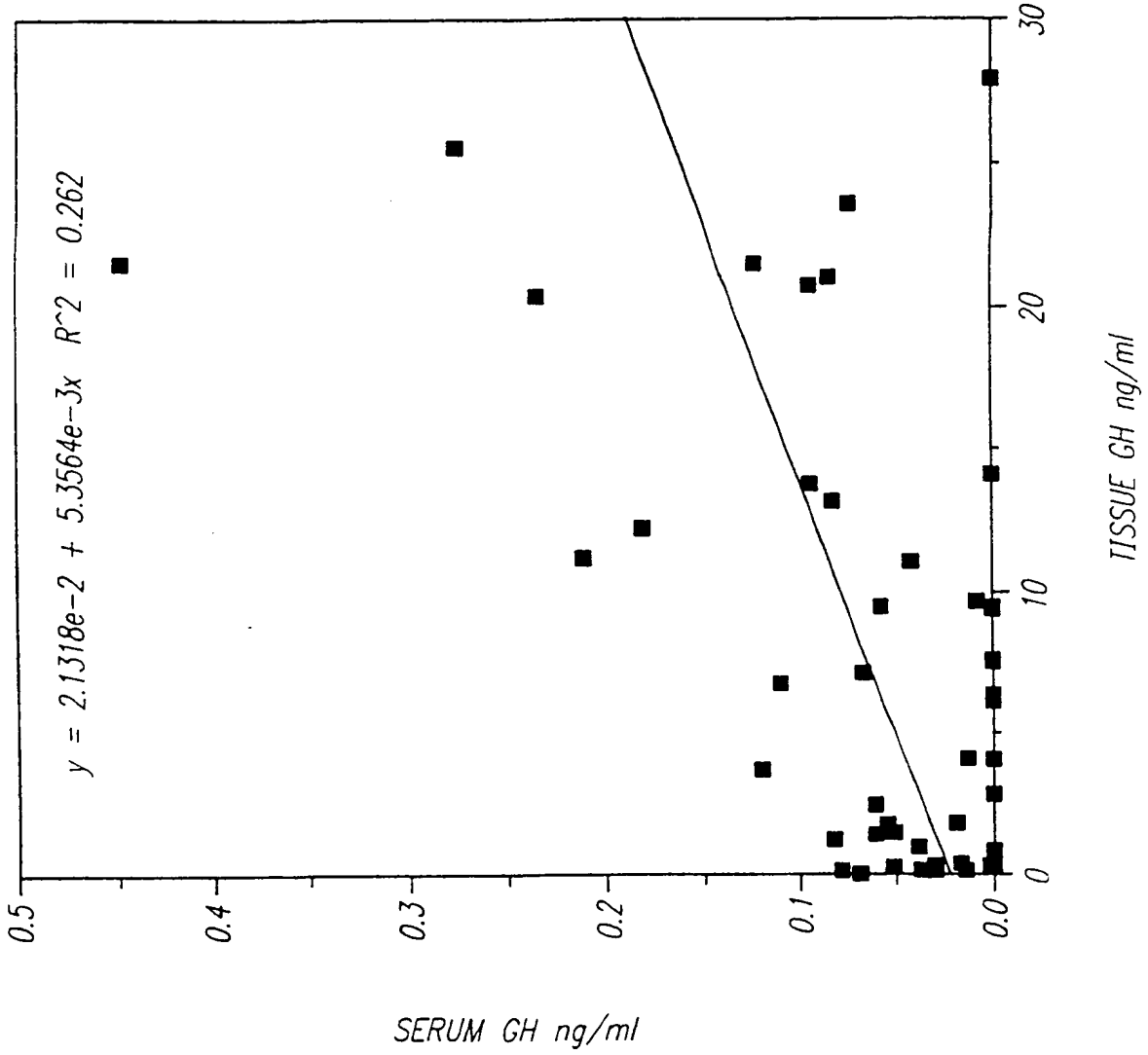


FIG. 6

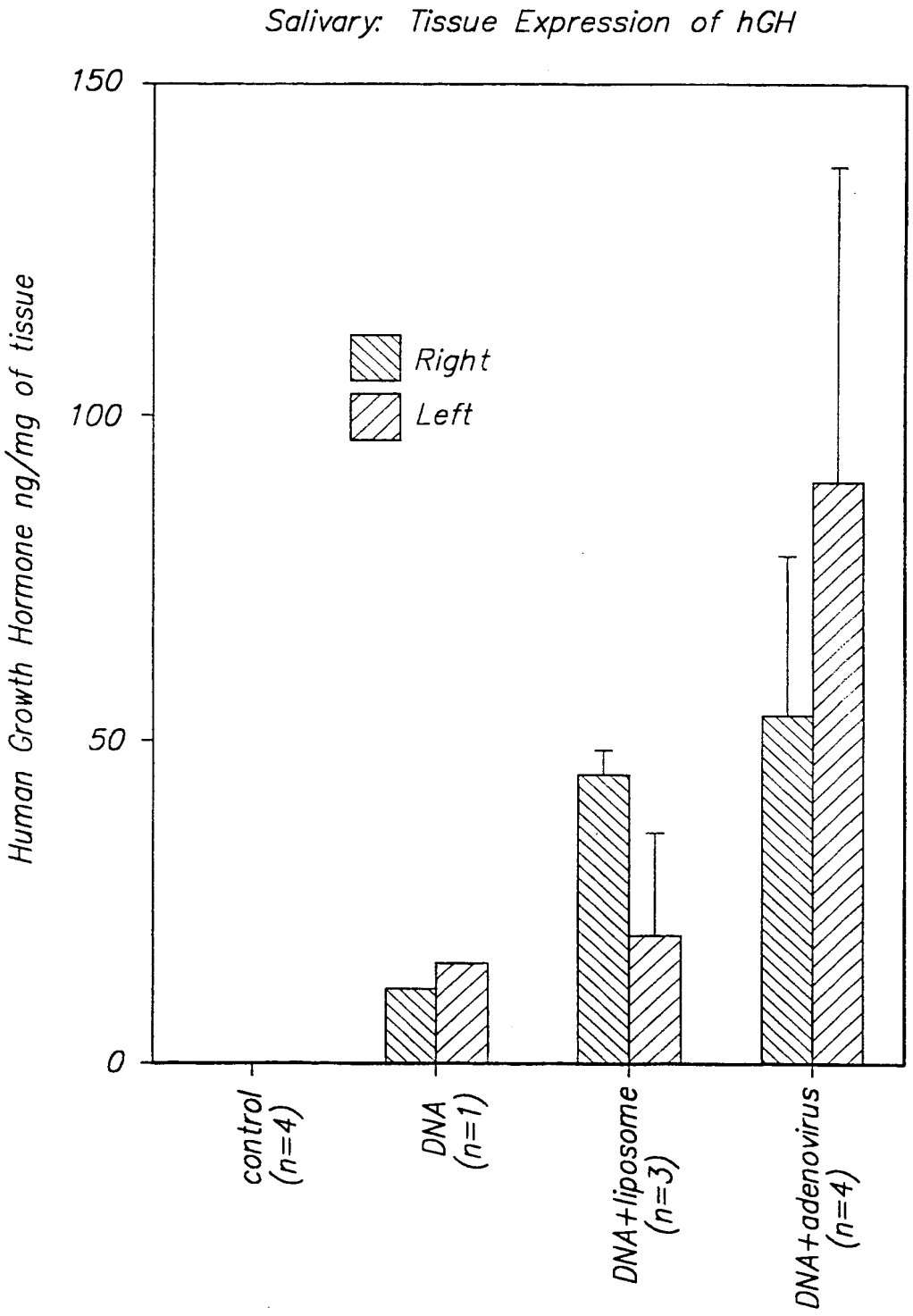


FIG. 7

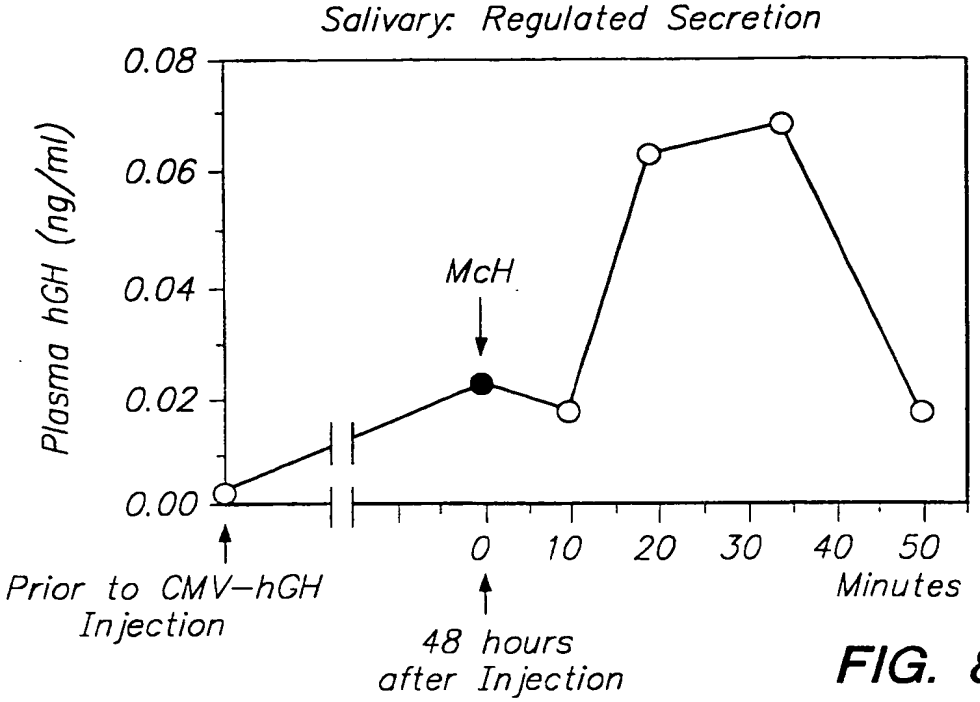


FIG. 8

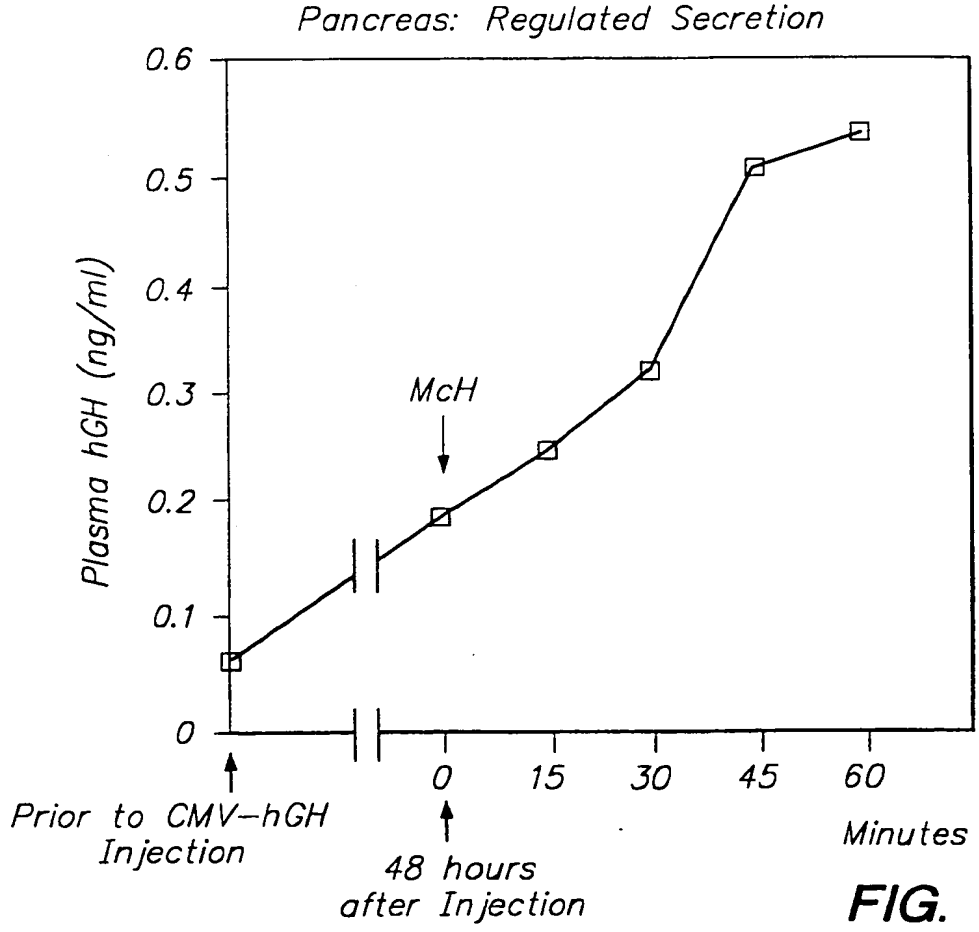


FIG. 9

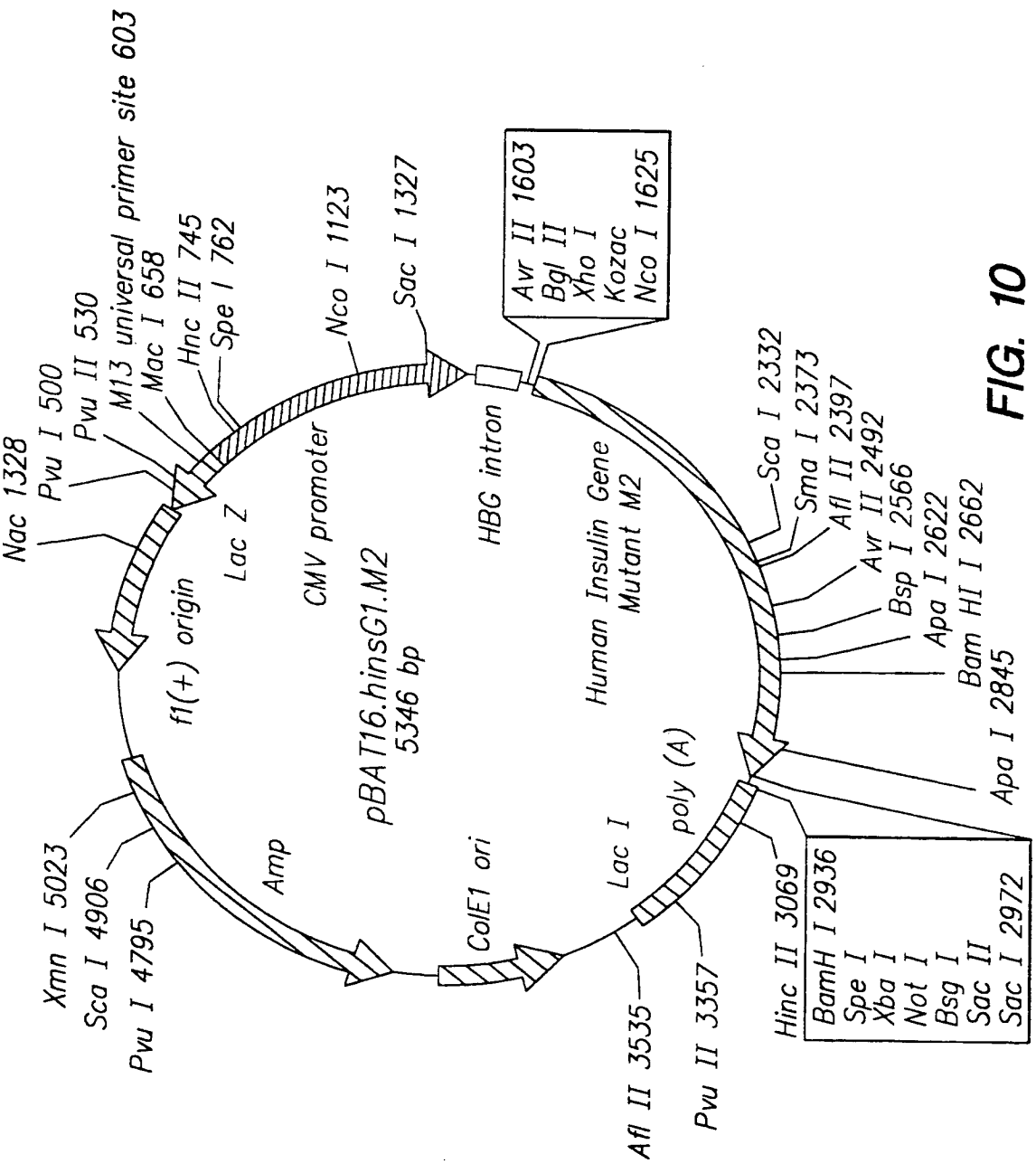


FIG. 10

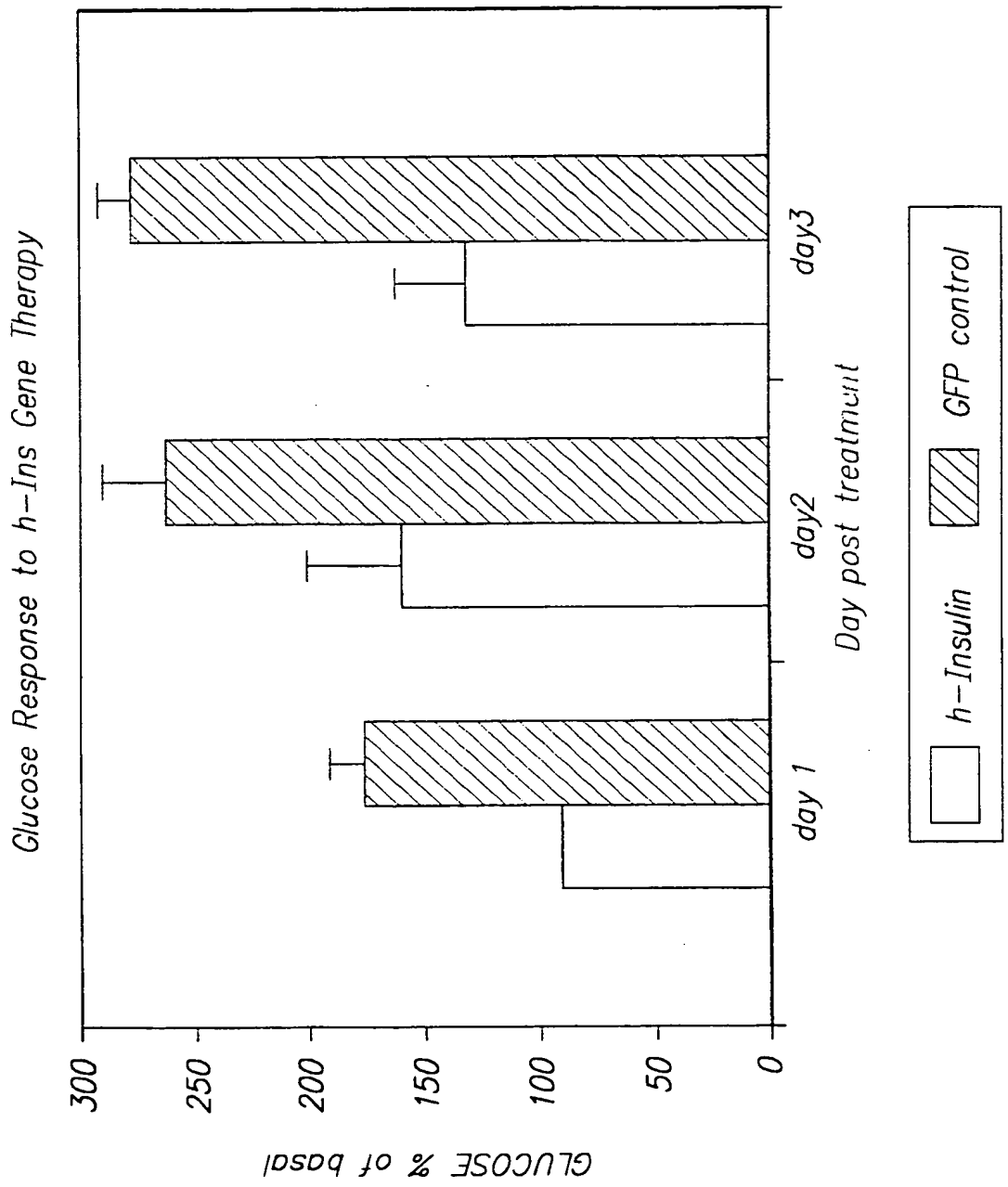

Glucose Response to h-Ins Gene Therapy
GLUCOSE % of basal
300
250
200
150
100
50
0
day 1
day2
day3
Day post treatment
h-Insulin
GFP control

FIG. 11

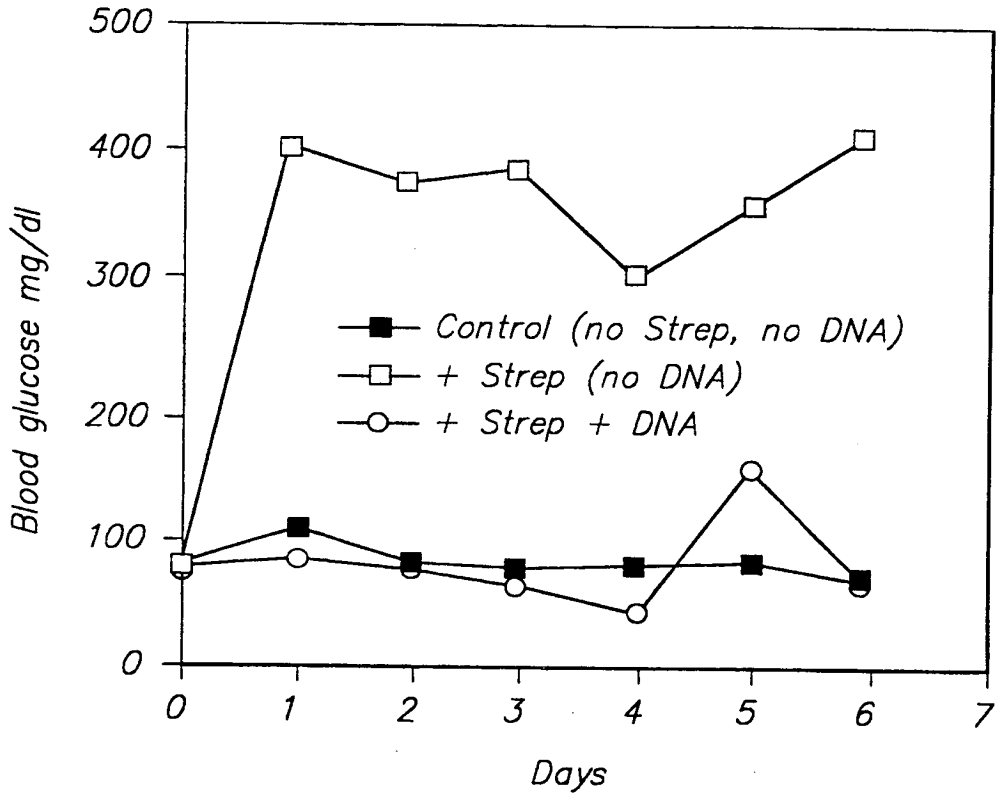


FIG. 12

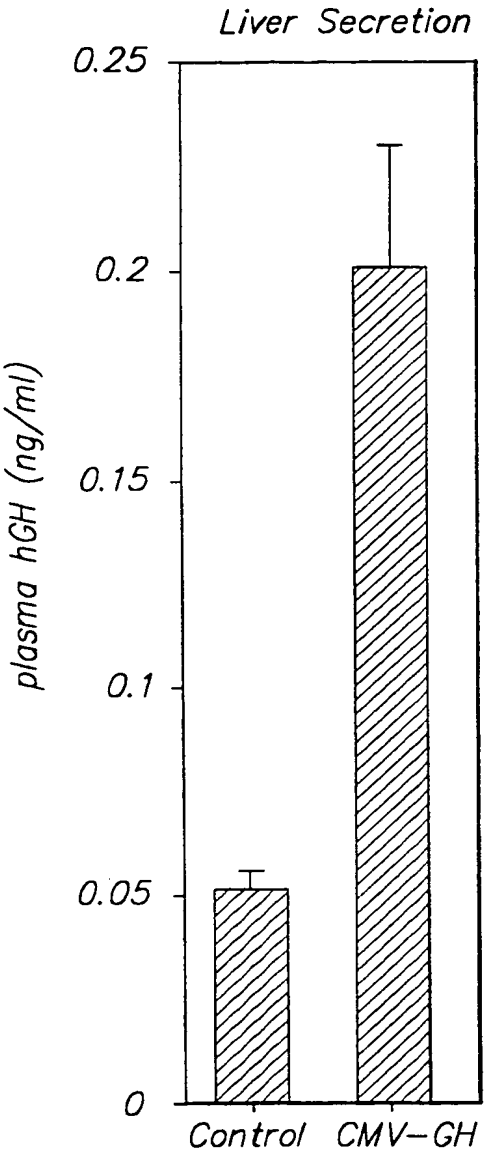

Liver Secretion
0.25
0.2
0.15
0.1
0.05
0
plasma hGH (ng/ml)
Control
CMV-GH

FIG. 13

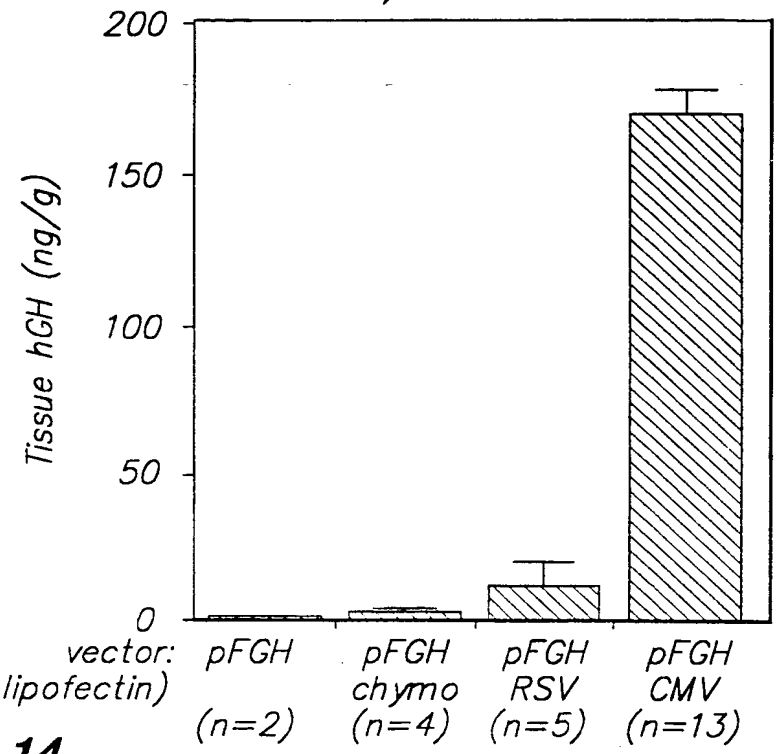


FIG. 14

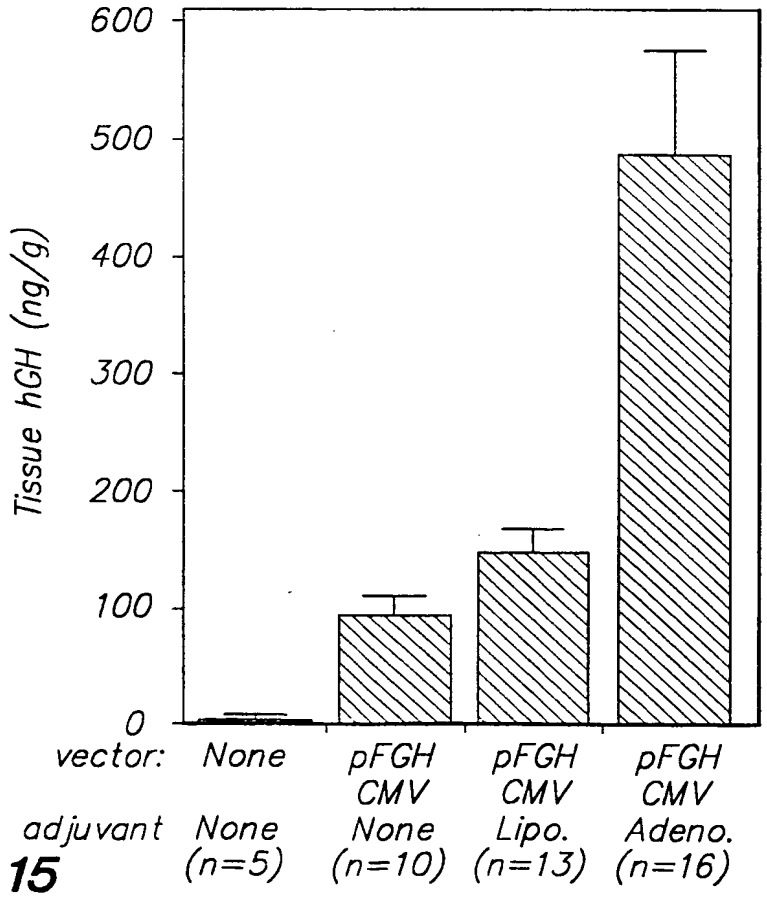


FIG. 15

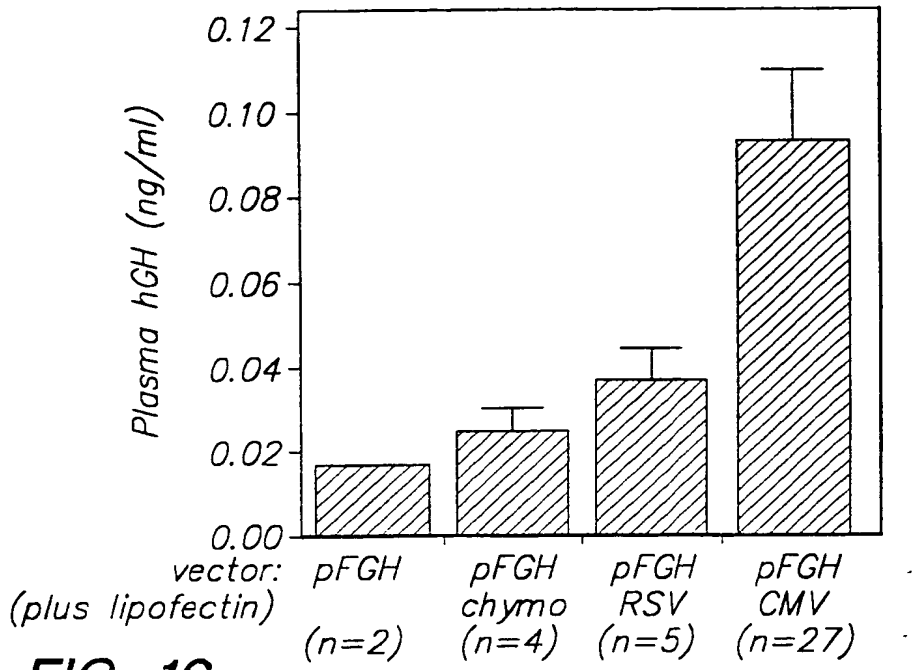


FIG. 16

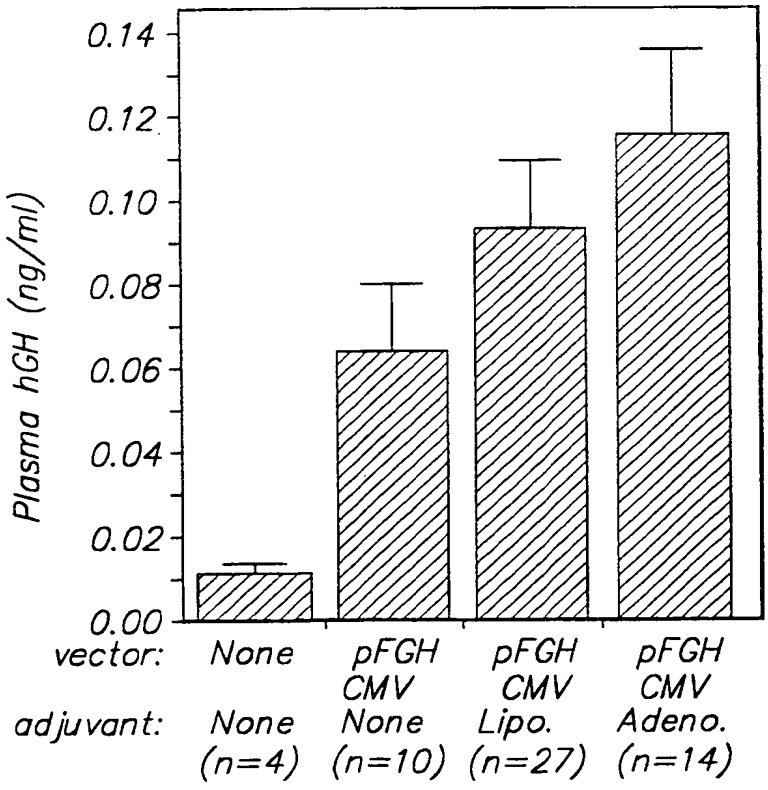


FIG. 17

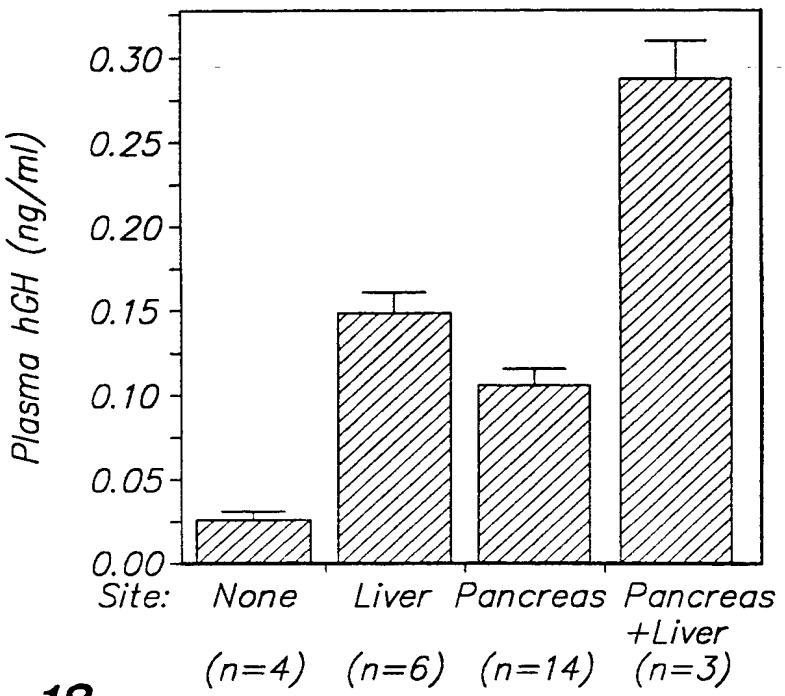


FIG. 18

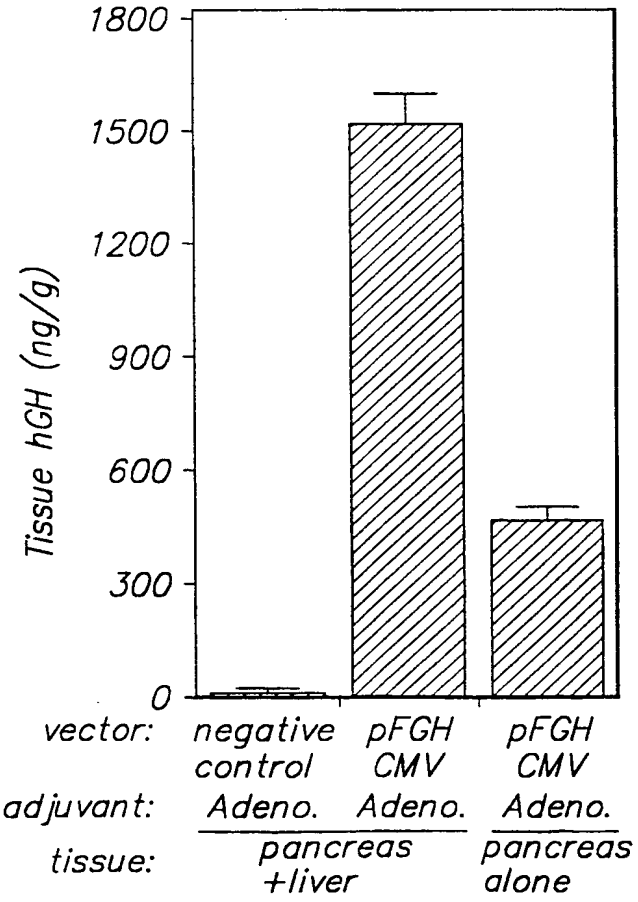


FIG. 19

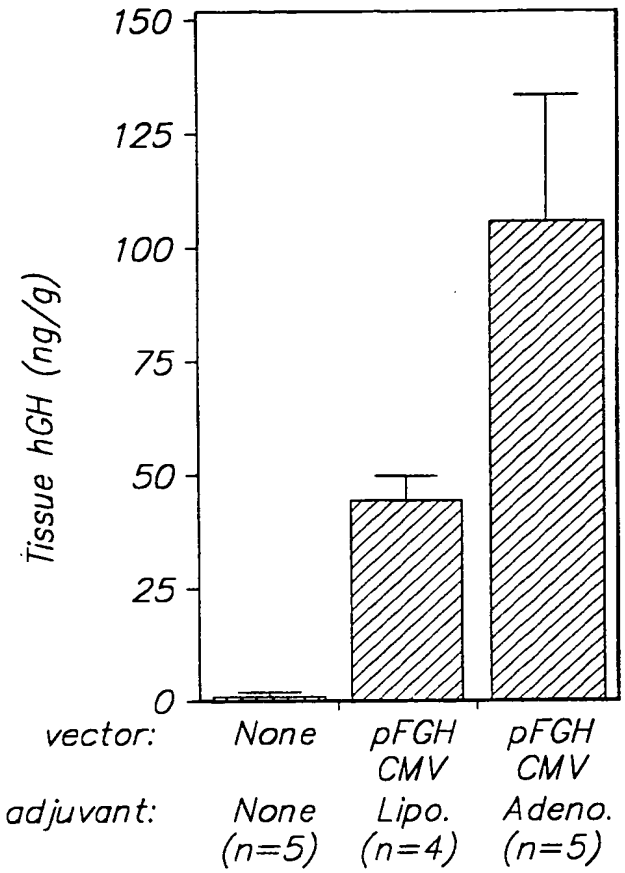


FIG. 20

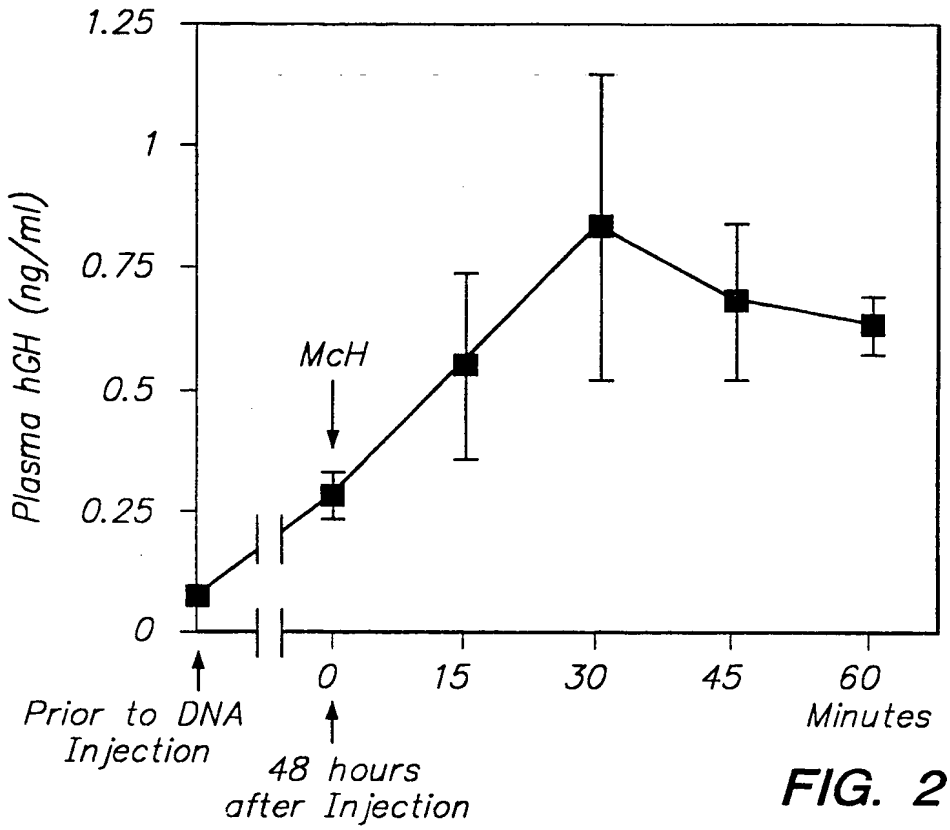


FIG. 21

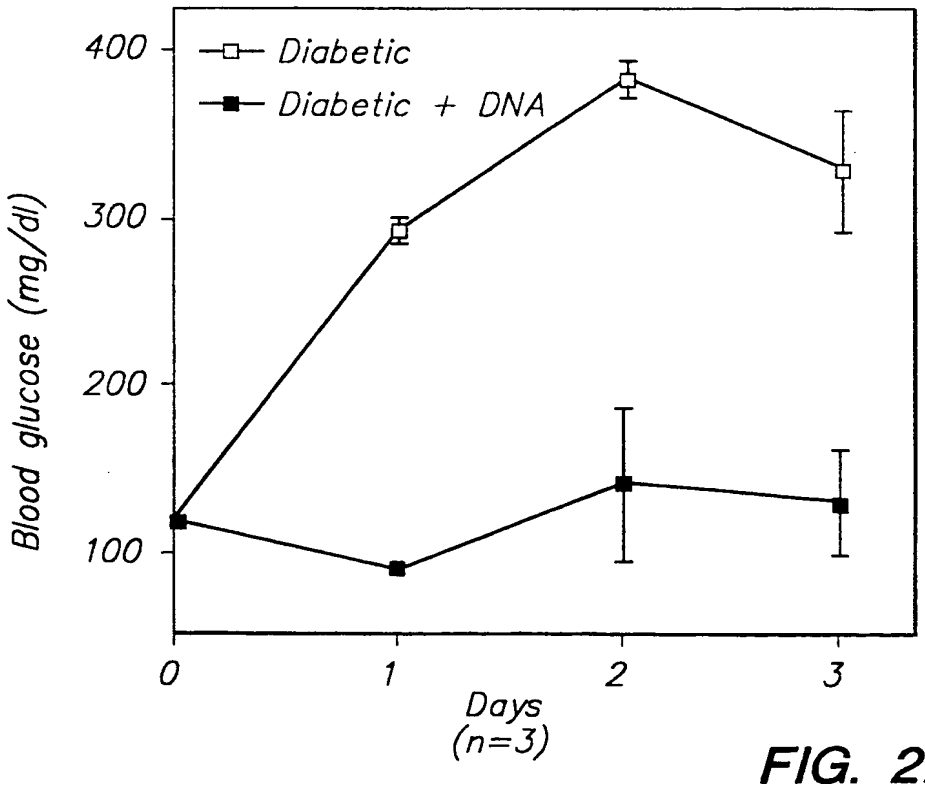


FIG. 22

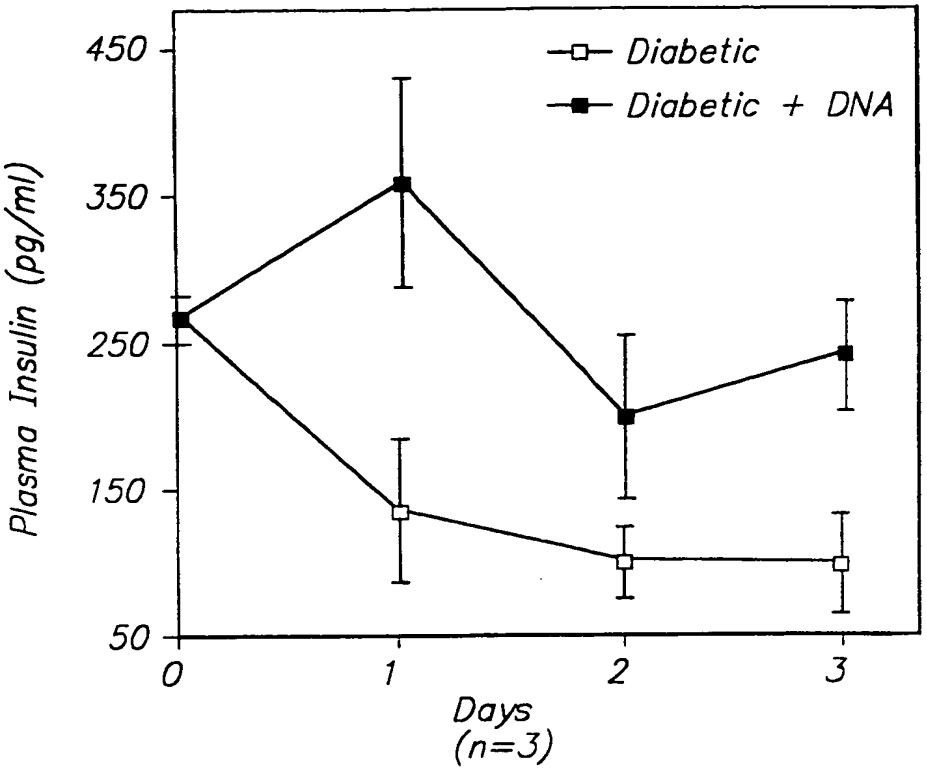


FIG. 23

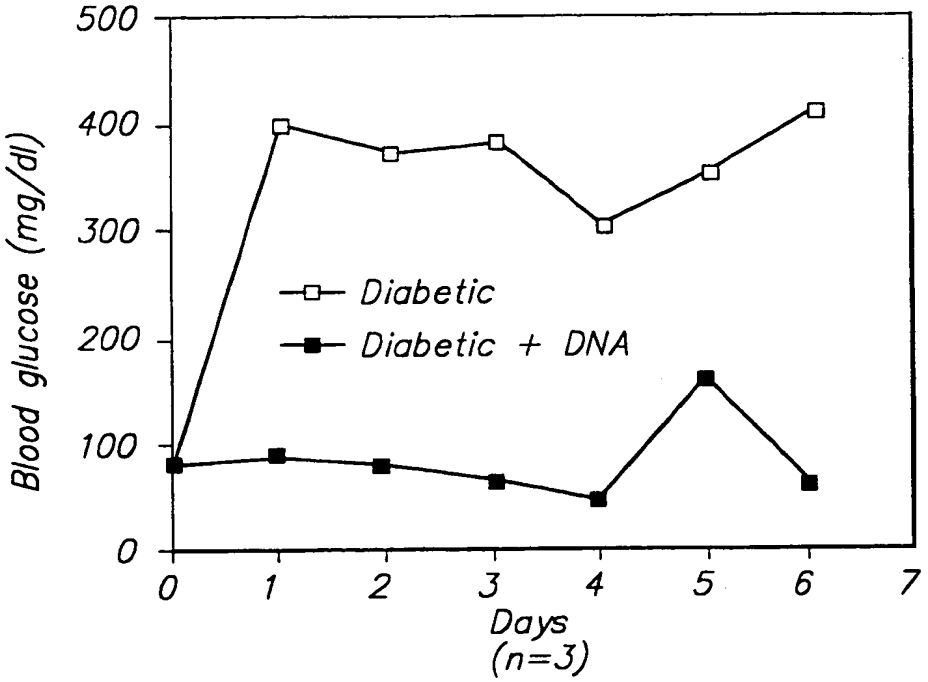


FIG. 24

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/20199

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :A01N 43/04, 63/00; C12N 13/00
US CL : 514/44; 424/93.21; 435/455

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 514/44; 424/93.21; 435/455

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, CAS ONLINE, MEDLINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	O'CONNELL et al. Transfer of a Gene encoding the Anticandidal Protein Histatin 3 to Salivary Glands. Human Gene Therapy. 01 December 1996, Vol. 7, pages 2255-2261, especially pages 2259 and 2260.	1-10
Y	FEHER et al. Exit Routes For Secretory Proteins From The Dog Pancreas. Acta Physiologica Academiae Scientiarum Hungaricae. 1980, Tomus 56, No. 4, pages 401-410, especially pages 407-409.	1-10
Y	KAGAMI et al. Evidence for the systemic Delivery of a Transgene Product from Salivary Glands. Human Gene Therapy. 10 November 1996, Vol. 7, pages 2177-2184, espcially pages 2181-2183.	1-10

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
18 DECEMBER 1998	25 JAN 1999
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer DAVE NGUYEN Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/20199

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	ADESANYA et al. Immediate Inflammatory responses to Adenovirus-Mediated Gene Transfer in rat Salivary Glands. Human Gene Therapy. 10 June 1996, Vol. 7, pages 1085-1093, abstract, espcially pages 1090 and 1091.	1-10
A,P	US 5,792,751 A (LEDLEY ET AL) 11 August 1998, columns 4-10.	1-10

PCT

LD INTELLECTUAL PROPERTY ORGANIZATIO
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
A23B 7/10, A23J 1/00, C07K 16/12, C07H 15/11, 21/02, 21/04, C12N 9/16, 15/85, 15/63, 1/20, C12Q 1/68

A1

(11) International Publication Number: WO 99/08539

(43) International Publication Date: 25 February 1999 (25.02.99)

(21) International Application Number: PCT/US98/16728

(22) International Filing Date: 13 August 1998 (13.08.98)

(30) Priority Data:
08/910,798 13 August 1997 (13.08.97) US

(71) Applicant: DIVERSA CORPORATION [US/US]; 10665 Sorrento Valley Road, San Diego, CA 92121 (US).

(72) Inventor: KRETZ, Keith; 973 Hawthorne Drive, San Marcos, CA 92069 (US).

(74) Agent: HAILE, Lisa, A.; Fish & Richardson P.C., Suite 1400, 4225 Executive Square, La Jolla, CA 92037 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: NOVEL PHYTASE

(57) Abstract

The invention provides a purified phytate enzyme derived from *Escherichia coli* B. The enzyme has a molecular weight of about 47.1 kilodaltons and has phytase activity (SEQ ID NO:2). The enzyme can be produced from native or recombinant host cells and can be used to aid in the digestion of phytate where desired. In particular, the phytase of the present invention can be used in animal feed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

NOVEL PHYTASE

Field of the Invention

This invention relates to newly identified polynucleotides, polypeptides encoded by such polynucleotides, the use of such polynucleotides and polypeptides, as well as the production and isolation of such polynucleotides and polypeptides. More particularly, the polypeptides of the present invention have been identified as phytases and in particular, enzymes having phytase activity.

Background

Minerals are essential elements for the growth of all organisms. For livestock production of monogastric animals (e.g., pigs, poultry) and fish, feed is commonly supplemented with minerals. Plant seeds are a rich source of minerals since they contain ions that are complexed with the phosphate groups of phytic acid. Ruminants do not require inorganic phosphate and minerals because microorganisms in the rumen produce enzymes that catalyze conversion of phytate (myo-inositol-hexaphosphate) to inositol and inorganic phosphate. In the process, minerals that have been complexed with phytate are released.

Phytate occurs as a source of stored phosphorous in virtually all plant feeds (Phytic Acid, Chemistry and Applications, E. Graf (Ed.), Pilatus Press: Minneapolis, MN, U.S.A., 1986). Phytic acid forms a normal part of the seed in cereals and legumes. It functions to bind dietary minerals that are essential to the new plant as it emerges from the seed. When the phosphate groups of phytic acid are removed by the seed enzyme phytase, the ability to bind metal ions is lost and the minerals become available to the plant. In livestock feed grains, the trace minerals bound by phytic acid are only partially available for absorption by monogastric animals, which lack phytase activity. Although some hydrolysis of phytate occurs in the colon, most phytate passes through the gastrointestinal tract of monogastric animals and is excreted in the manure contributing

to fecal phosphate pollution problems in areas of intense livestock production. Inorganic phosphorous released in the colon has no nutritional value to livestock because inorganic phosphorous is absorbed only in the small intestine. Thus, a significant amount of the nutritionally important dietary minerals are potentially not available to monogastric animals.

Conversion of phytate to inositol and inorganic phosphorous can be catalyzed by microbial enzymes referred to broadly as phytases. Phytases such as phytase #EC 3.1.3.8 are capable of catalyzing hydrolysis of myo-inositol hexaphosphate to D-myo-inositol 1,2,4,5,6-pentaphosphate and orthophosphate. Certain fungal phytases reportedly hydrolyze inositol pentaphosphate to tetra-, tri-, and lower phosphates; e.g., A. ficuum phytases reportedly produce mixtures of myoinositol di- and mono-phosphate (Ullah, 1988). Phytase producing microorganisms comprise bacteria such as Bacillus subtilis (V.K. Powar and V.J. Jagannathan, J. Bacteriol. 151:1102-1108, 1982) and Pseudomonas (D.J. Cosgrove, Austral. J. Biol. Sci. 2:1207-1220, 1970); yeasts such as Sacchoromyces cerevisiae (N.R. Nayini and P. Markakis, Lebensmittel Wissenschaft und Technologie 17:24-26, 1984); and fungi such as Aspergillus terreus (K. Yamada, et al., Agric. Biol Chem. 32:1275-1282, 1968). The possible use of microbes capable of producing phytase as a feed additive for monogastric animals has been reported previously (Shieh and Ware, U.S. Patent No. 3,297,548; Nelson, T.S. et al., J. Nutrition 101:1289-1294, 1971).

Microbial phytases may also reportedly be useful for producing animal feed from certain industrial processes, e.g., wheat and corn waste products. The wet milling process of corn produces glutens sold as animal feeds. Addition of phytase may reportedly improve the nutritional value of the feed product. Fungal phytase enzymes and process conditions (t~50(C and pH ~5.5) have been reported previously in European Patent Application 0 321 004. In processing soybean meal the presence of phytate reportedly renders the meal and wastes unsuitable for feeds used in rearing fish, poultry and other non-ruminants as well as calves fed on milk. Phytase is reportedly useful for improving the nutrient and

commercial value of this high protein soy material (see Finase Enzymes by Alko, Rajamaki, Finland). A combination of phytase and a pH 2.5 optimum acid phosphatase form A. niger has been used by Alko, Ltd as an animal feed supplement in their phytic acid degradative product Finas F and Finase S. A cost-effective source of phytase would greatly enhance the value of soybean meals as an animal feed (Shieh et al., 1969).

Phytase and less specific acid phosphatases are produced by the fungus Aspergillus ficuum as extracellular enzymes (Shieh et al., 1969). Ullah reportedly purified a phytase from wild-type A. ficuum that had an apparent molecular weight of 61.7 kDA (on SDS-PAGE; as corrected for glycosylation); pH optima at pH 2.5 and pH 5.5; a Km of about 40(m; and, a specific activity of about 50U/mg (Ullah, A., Preparative Biochem 18:443-458, 1988); PCT patent application WO 91/05053 also reportedly discloses isolation and molecular cloning of a phytase from Aspergillus ficuum with pH optima at pH 2.5 and pH 5.5, a Km of about 250(m, and specific activity of about 100U/mg protein.

Acid phosphatases are enzymes that catalytically hydrolyze a wide variety of phosphate esters and usually exhibit pH optima below 6.0 (Hollander, 1971); e.g., #EC 3.1.3.2 catalyzes hydrolysis of orthophosphoric monoesters to orthophosphate products. An acid phosphatase has reportedly been purified from A. ficuum. The deglycosylated form of the acid phosphatase has an apparent molecular weight of 32.6 kDa (Ullah et al., 1987).

The object of the present invention provides a recombinant phytase isolated from *Escherichia coli* B that improves the efficiency of release of phosphorous from phytate and the salts of phytic acid. Another object of the present invention provides a source of a recombinant enzyme that is suitable for commercial use in feeds and industrial processes with minimal processing.

Summary of the Invention

The present invention provides a polynucleotide and a polypeptide encoded thereby which has been identified as a phytase enzyme having phytase activity. In accordance with one aspect of the present invention, there is provided a novel recombinant enzyme, as well as active fragments, analogs and derivatives thereof.

In accordance with another aspect of the present invention, there are provided isolated nucleic acid molecules encoding the enzyme of the present invention including mRNA, DNA, cDNA, genomic DNA as well as active analogs and fragments of such enzyme.

In accordance with yet a further aspect of the present invention, there is provided a process for producing such polypeptide by recombinant techniques comprising culturing recombinant prokaryotic and/or eukaryotic host cells, containing a nucleic acid sequence encoding an enzyme of the present invention, under conditions promoting expression of said enzyme and subsequent recovery of said enzyme.

In accordance with yet a further aspect of the present invention, there is provided a process for utilizing such enzyme, or polynucleotide encoding such enzymes for use in commercial processes, such as, for example, processes that liberate minerals from phytates in plant materials either in vitro, i.e., in feed treatment processes, or in vivo, i.e., by administering the enzyme to animals.

In accordance with yet a further aspect of the present invention, there is provided a process for utilizing such enzymes, or polynucleotides encoding such enzymes, for *in vitro* purposes related to scientific research, for example, to generate probes for identifying similar sequences which might encode similar enzymes from other organisms.

These and other aspects of the present invention should be apparent to those skilled in the art from the teachings herein.

BRIEF DESCRIPTION OF THE DRAWINGS

The following drawings are illustrative of embodiments of the invention and are not meant to limit the scope of the invention as encompassed by the claims.

Figure 1 shows the nucleotide and deduced amino acid sequences the enzyme of the present invention. Sequencing was performed using a 378 automated DNA sequencer (Applied Biosystems, Inc.).

Figure 2 shows the pH and temperature profile and stability data for the phytase enzyme of the present invention. The assay used for these analysis is the following for the detection of phytase activity: Phytase activity is measured by incubating 150(l of the enzyme preparation with 600(l of 2 mM sodium phytate in 100 mM Tris HCl buffer pH 7.5, supplemented with 1mM $CaCl_2$ for 30 minutes at 37(C. After incubation the reaction is stopped by adding 750(l of 5% trichloroacetic acid. Phosphate released was measured against phosphate standard spectrophotometrically at 700nm after adding 1500(l of the color reagent (4 volumes of 1.5% ammonium molybdate in 5.5% sulfuric acid and 1 volume of 2.7% ferrous sulfate; Shimizu, M., 1992; Biosci. Biotech. Biochem., 56:1266-1269). OD at 700nm is indicated on the Y-axis of the graphs in Figure 2. Temperature or pH is indicated on the X-axis of the graphs.

DETAILED DESCRIPTION OF THE INVENTION

The term "gene" means the segment of DNA involved in producing a polypeptide chain; it includes regions preceding and following the coding region (leader and trailer) as well as intervening sequences (introns) between individual coding segments (exons).

A coding sequence is "operably linked to" another coding sequence when RNA polymerase will transcribe the two coding sequences into a single mRNA, which is then translated into a single polypeptide having amino acids derived from both coding sequences. The coding sequences need not be contiguous to one another so long as the expressed sequences are ultimately processed to produce the desired protein.

"Recombinant" enzymes refer to enzymes produced by recombinant DNA techniques; *i.e.*, produced from cells transformed by an exogenous DNA construct encoding the desired enzyme. "Synthetic" enzymes are those prepared by chemical synthesis.

A DNA "coding sequence of" or a "nucleotide sequence encoding" a particular enzyme, is a DNA sequence which is transcribed and translated into an enzyme when placed under the control of appropriate regulatory sequences. A "promotor sequence" is a DNA regulatory region capable of binding RNA polymerase in a cell and initiating transcription of a downstream (3' direction) coding sequence. The promoter is part of the DNA sequence. This sequence region has a start codon at its 3' terminus. The promoter sequence does include the minimum number of bases where elements necessary to initiate transcription at levels detectable above background. However, after the RNA polymerase binds the sequence and transcription is initiated at the start codon (3' terminus with a promoter), transcription proceeds downstream in the 3' direction. Within the promotor sequence will be found a transcription initiation site (conveniently defined by mapping with nuclease S1) as well as protein binding domains (consensus sequences) responsible for the binding of RNA polymerase.

The present invention provides purified a recombinant enzyme that catalyzes the hydrolysis of phytate to inositol and free phosphate with release of minerals from the phytic acid complex. An exemplary purified enzyme is a phytase derived from *Escherichia coli* B. This exemplary enzyme is shown in Figure 1, SEQ ID NO:2.

The polynucleotide encoding SEQ ID NO:2 was originally recovered from genomic DNA isolated from *Escherichia coli* B as described below. It contains an open reading frame encoding a protein of 432 amino acid residues.

In one embodiment, the phytase enzyme of SEQ ID NO:2 of the present invention has a molecular weight of about 47,056 kilodaltons as measured by SDS-PAGE gel electrophoresis and an inferred molecular weight from the nucleotide sequence of the gene. The pI is 6.70. The pH and temperature profile and stability data for this enzyme is presented in Figure 2. This purified enzyme may be used to catalyze the hydrolysis of phytate to inositol and free phosphate where desired. The phytase enzyme of the present invention has a high thermostability.

In accordance with an aspect of the present invention, there are provided isolated nucleic acid molecules (polynucleotides) which encode for the mature enzyme having the deduced amino acid sequence of Figure 1.

This invention can be used to isolate nucleic acid sequences substantially similar to the isolated nucleic acid molecule encoding an phytase enzyme disclosed in Figure 1 (SEQ ID NO:1),. Isolated nucleic acid sequences are substantially similar if: (i) they are capable of hybridizing under stringent conditions, hereinafter described, to SEQ ID NO:1; or (ii) they encode DNA sequences which are degenerate to SEQ ID NO:1. Degenerate DNA sequences encode the amino acid sequence of SEQ ID NO:2, but have variations in the nucleotide coding sequences. As used herein, "substantially similar" refers to the sequences having similar identity to the sequences of the instant invention.

The nucleotide sequences that are substantially similar can be identified by hybridization or by sequence comparison. Enzyme sequences that are substantially similar can be identified by one or more of the following: proteolytic digestion, gel electrophoresis and/or microsequencing.

One means for isolating a nucleic acid molecule encoding a phytase enzyme is to probe a genomic gene library with a natural or artificially designed probe using art recognized procedures (see, for example: Current Protocols in Molecular Biology, Ausubel F.M. *et al.* (EDS.) Green Publishing Company Assoc. and John Wiley Interscience, New York, 1989, 1992). It is appreciated to one skilled in the art that SEQ ID NO:1, or fragments thereof (comprising at least 15 contiguous nucleotides), is a particularly useful probe. Other particular useful probes for this purpose are hybridizable fragments to the sequences of SEQ ID NO:1 (comprising at least 10 contiguous nucleotides and at least 70% complementary to a target sequence), is a particularly useful probe. Other particular useful probes for this purpose are hybridizable fragments to the sequences of SEQ ID NO:1 (*i.e.*, comprising at least 10 contiguous nucleotides and at least 70% complementary to a target sequence).

With respect to nucleic acid sequences which hybridize to specific nucleic acid sequences disclosed herein, hybridization may be carried out under conditions of reduced stringency, medium stringency or even stringent conditions. As an example of oligonucleotide hybridization, a polymer membrane containing immobilized denatured nucleic acid is first prehybridized for 30 minutes at 45(C in a solution consisting of 0.9 M NaCl, 50 mM NaH_2PO_4, pH 7.0, 5.0 mM Na_2EDTA, 0.5% SDS, 10X Denhardt's, and 0.5 mg/mL polyriboadenylic acid. Approximately 2 X 10^7 cpm (specific activity 4-9 X 10^8 cpm/ug) of^{32} P end-labeled oligonucleotide probe are then added to the solution. After 12-16 hours of incubation, the membrane is washed for 30 minutes at room temperature in 1X SET (150 mM NaCl, 20 mM Tris hydrochloride, pH 7.8, 1 mM Na_2EDTA) containing 0.5% SDS, followed by a 30 minute wash in fresh 1X SET at

Tm-10(C for the oligo-nucleotide probe. The membrane is then exposed to auto-radiographic film for detection of hybridization signals.

Stringent conditions means hybridization will occur only if there is at least 90% identity, preferably at least 95% identity and most preferably at least 97% identity between the sequences. *See* J. Sambrook *et al.*, Molecular Cloning, A Laboratory Manual (2d Ed. 1989) (Cold Spring Harbor Laboratory) which is hereby incorporated by reference in its entirety.

"Identity" as the term is used herein, refers to a polynucleotide sequence which comprises a percentage of the same bases as a reference polynucleotide (SEQ ID NO:1). For example, a polynucleotide which is at least 90% identical to a reference polynucleotide, has polynucleotide bases which are identical in 90% of the bases which make up the reference polynucleotide and may have different bases in 10% of the bases which comprise that polynucleotide sequence.

The present invention also relates to polynucleotides which differ from the reference polynucleotide such that the changes are silent changes, for example the changes do not alter the amino acid sequence encoded by the polynucleotide. The present invention also relates to nucleotide changes which result in amino acid substitutions, additions, deletions, fusions and truncations in the enzyme encoded by the reference polynucleotide (SEQ ID NO:1). In a preferred aspect of the invention these enzymes retain the same biological action as the enzyme encoded by the reference polynucleotide.

It is also appreciated that such probes can be and are preferably labeled with an analytically detectable reagent to facilitate identification of the probe. Useful reagents include but are not limited to radioactivity, fluorescent dyes or enzymes capable of catalyzing the formation of a detectable product. The probes are thus useful to isolate complementary copies of DNA from other animal sources or to screen such sources for

related sequences.

The present invention provides a substantially pure phytase enzyme. The term "substantially pure" is used herein to describe a molecule, such as a polypeptide (*e.g.*, a phytase polypeptide, or a fragment thereof) that is substantially free of other proteins, lipids, carbohydrates, nucleic acids, and other biological materials with which it is naturally associated. For example, a substantially pure molecule, such as a polypeptide, can be at least 60%, by dry weight, the molecule of interest. The purity of the polypeptides can be determined using standard methods including, *e.g.*, polyacrylamide gel electrophoresis (*e.g.*, SDS-PAGE), column chromatography (*e.g.*, high performance liquid chromatography (HPLC)), and amino- terminal amino acid sequence analysis.

The phytase polypeptide included in the invention can have the amino acid sequences of Phytase shown in Figure 1 (SEQ ID NO:1). Phytase polypeptides, such as those isolated from *E.coli* B , can be characterized by catalyzing the hydrolysis of phytate to inositol and free phosphate with the release of minerals from the phytic acid complex.

Also included in the invention are polypeptides having sequences that are "substantially identical" to the sequence of a phytase polypeptide, such as one of SEQ ID 1. A "substantially identical" amino acid sequence is a sequence that differs from a reference sequence only by conservative amino acid substitutions, for example, substitutions of one amino acid for another of the same class (*e.g.*, substitution of one hydrophobic amino acid, such as isoleucine, valine, leucine, or methionine, for another, or substitution of one polar amino acid for another, such as substitution of arginine for lysine, glutamic acid for aspartic acid, or glutamine for asparagine).

Fragments of the phytase polypeptide of the present invention can retain at least one phytase -specific activity or epitope. Phytase activity can be assayed by examining the

catalysis of phytate to inositol and free phosphate. For example, a phytase polypeptide fragment containing, *e.g.*, at least 8-10 amino acids can be used as an immunogen in the production of phytase-specific antibodies. The fragment can contain, for example, an amino acid sequence that is conserved in phytases, and this amino acid sequence can contain amino acids that are conserved in phytases. Such fragments can easily be identified by comparing the sequences of phytases found in Figure 1. In addition to their use as peptide immunogens, the above-described phytase fragments can be used in immunoassays, such as ELISAs, to detect the presence of phytase-specific antibodies in samples.

Other phytase polypeptides included in the invention are polypeptides having amino acid sequences that are at least 50% identical to the amino acid sequence of a phytase polypeptide, such as SEQ ID NO:2. The length of comparison in determining amino acid sequence homology can be, for example, at least 15 amino acids, for example, at least 20, 25, or 35 amino acids. Homology can be measured using standard sequence analysis software (e.g., Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705; also see Ausubel, et al., supra). Such procedures and algorithms include, for example, a BLAST program (Basic Local Alignment Search Tool at the National Center for Biological Information), ALIGN, AMAS (Analysis of Multiply Aligned Sequences), AMPS (Protein Multiple Sequence Alignment), ASSET (Aligned Segment Statistical Evaluation Tool), BANDS, BESTSCOR, BIOSCAN (Biological Sequence Comparative Analysis Node), BLIMPS (BLocks IMProved Searcher), FASTA, Intervals & Points, BMB, CLUSTAL V, CLUSTAL W, CONSENSUS, LCONSENSUS, WCONSENSUS, Smith-Waterman algorithm, DARWIN, Las Vegas algorithm, FNAT (Forced Nucleotide Alignment Tool), Framealign, Framesearch, DYNAMIC, FILTER, FSAP (Fristensky Sequence Analysis Package), GAP (Global Alignment Program), GENAL, GIBBS, GenQuest, ISSC (Sensitive Sequence Comparison), LALIGN (Local Sequence Alignment), LCP (Local Content Program), MACAW (Multiple Alignment

Construction & Analysis Workbench), MAP (Multiple Alignment Program), MBLKP, MBLKN, PIMA (Pattern-Induced Multi-sequence Alignment), SAGA (Sequence Alignmnet by Genetic ALgorithm) and WHAT-IF.

The phytase polypeptide of the invention can be obtained using any of several standard methods. For example, phytase polypeptides can be produced in a standard recombinant expression systems (see below), chemically synthesized (this approach may be limited to small phytase peptide fragments), or purified from organisms in which they are naturally expressed.

The invention also provides isolated nucleic acid molecules that encode the phytase polypeptide described above. For example, nucleic acids that encode SEQ ID NO:1 are included in the invention. These nucleic acids can contain naturally occurring nucleotide sequences, or sequences that differ from those of the naturally occurring nucleic acids that encode phytases, but encode the same amino acids, due to the degeneracy of the genetic code. The nucleic acids of the invention can contain DNA or RNA nucleotides, or combinations or modifications thereof. Exemplary nucleic acids of the invention are shown in SEQ ID NO:1.

By "isolated nucleic acid" is meant a nucleic acid, *e.g.*, a DNA or RNA molecule, that is not immediately contiguous with the 5' and 3' flanking sequences with which it normally is immediately contiguous when present in the naturally occurring genome of the organism from which it is derived. The term thus describes, for example, a nucleic acid that is incorporated into a vector, such as a plasmid or viral vector; a nucleic acid that is incorporated into the genome of a heterologous cell (or the genome of a homologous cell, but at a site different from that at which it naturally occurs); and a nucleic acid that exists as a separate molecule, *e.g.*, a DNA fragment produced by PCR amplification or restriction enzyme digestion, or an RNA molecule produced by *in vitro* transcription. The term also describes a recombinant nucleic acid that forms part of a

hybrid gene encoding additional polypeptide sequences that can be used, for example, in the production of a fusion protein.

The nucleic acid molecules of the invention can be used as templates in standard methods for production of phytase gene products (*e.g.*, phytase RNAs and phytase polypeptides). In addition, the nucleic acid molecules that encode phytase polypeptides (and fragments thereof) and related nucleic acids, such as (1) nucleic acids containing sequences that are complementary to, or that hybridize to, nucleic acids encoding phytase polypeptides, or fragments thereof (*e.g.*, fragments containing at least 10, 12, 15, 20, or 25 nucleotides); and (2) nucleic acids containing sequences that hybridize to sequences that are complementary to nucleic acids encoding phytase polypeptides, or fragments thereof (*e.g.*, fragments containing at least 10, 12, 15, 20, or 25 nucleotides); can be used in methods focused on their hybridization properties. For example, as is described in further detail below, such nucleic acid molecules can be used in the following methods: PCR methods for synthesizing phytase nucleic acids, methods for detecting the presence of a phytase nucleic acid in a sample, screening methods for identifying nucleic acids encoding new phytase family members. Oligonucleotide probes useful for screening methods are from 10 to about 150 nucleotides in length. Further, such probes are preferably 10 to about 100 nucleotides in length and more preferably from 10 to about 50 nucleotides in length.

The invention also includes methods for identifying nucleic acid molecules that encode members of the phytase polypeptide family in addition to SEQ ID NO:1. In these methods, a sample, *e.g.*, a nucleic acid library, such as a cDNA library, that contains a nucleic acid encoding a phytase polypeptide is screened with a phytase-specific probe, *e.g.*, a phytase-specific nucleic acid probe. Phytase-specific nucleic acid probes are nucleic acid molecules (*e.g.*, molecules containing DNA or RNA nucleotides, or combinations or modifications thereof) that specifically hybridize to nucleic acids encoding phytase polypeptides, or to complementary sequences thereof. The term

"phytase-specific probe," in the context of this method of invention, refers to probes that bind to nucleic acids encoding phytase polypeptides, or to complementary sequences thereof, to a detectably greater extent than to nucleic acids encoding other enzymes, or to complementary sequences thereof.

The invention facilitates production of phytase-specific nucleic acid probes. Methods for obtaining such probes can be designed based on the amino acid sequences shown in Figure 1. The probes, which can contain at least 10, *e.g.*,at least 15, 25, 35, 50, 100, or 150 nucleotides, can be produced using any of several standard methods (see, *e.g.*, Ausubel, *et al., supra*). For example, preferably, the probes are generated using PCR amplification methods. In these methods, primers are designed that correspond to phytase-conserved sequences (see Figure 1), which can include phytase-specific amino acids, and the resulting PCR product is used as a probe to screen a nucleic acid library, such as a cDNA library.

The coding sequences for the phytase enzymes of the present invention were identified by preparing *E.coli* B genomic DNA, for example, and recovering (via, for example, PCR amplification) from the genomic DNA, DNA encoding phytase activity. Such methods for recovery are well-known in the art. One means, for example, comprises designing amplification primers to recover the coding sequence, amplifying the gene from the genomic DNA, subcloning the DNA into a vector, transforming the resulting construct into a host strain, and expressing the phytase enzyme for evaluation. Such procedures are well known in the art and methods are provided, for example, in Maniatis, T., *et al.*, Molecular Cloning, Cold Spring Harbor Press, New York, 1982, which is hereby incorporated by reference in its entirety.

In a preferred embodiment, the enzyme of the present invention, was isolated from an *E.coli* B genomic DNA by the following technique:

E.coli B genomic DNA was obtained from Sigma (Catalog # D-2001), St. Louis, New

Jersey.

The following primers were used to amplify the gene directly from the genomic DNA:

5' primer gtttctgaattcaaggaggaatttaaATGAAAGCGATCTTAATCCCATT

3' primer gtttctggatccTTACAAACTGCACGCCGGTAT

Pfu polymerase was used according to manufacturers protocol (Stratagene Cloning Systems, Inc., La Jolla, CA).

PCR product and pQE60 vector (Qiagen) were both digested with EcoRI and BglII restriction endonucleases (New England Biolabs) according to manufacturers protocols. Ligation and transformation into, and expression in M15 pREP4 host cells (Qiagen) yields c-term 6X-His tagged protein.

Fragments of the full length gene of the present invention may be used as a hybridization probe for a cDNA or a genomic library to isolate the full length DNA and to isolate other DNAs which have a high sequence similarity to the gene or similar biological activity. Probes of this type have at least 10, preferably at least 15, and even more preferably at least 30 bases and may contain, for example, at least 50 or more bases. The probe may also be used to identify a DNA clone corresponding to a full length transcript and a genomic clone or clones that contain the complete gene including regulatory and promotor regions, exons, and introns.

The isolated nucleic acid sequences and other enzymes may then be measured for retention of biological activity characteristic to the enzyme of the present invention, for example, in an assay for detecting enzymatic phytase activity. Such enzymes include truncated forms of phytase, and variants such as deletion and insertion variants.

Examples of such assays include the following assay for the detection of phytase activity: Phytase activity can be measured by incubating 150(l of the enzyme preparation with 600(l of 2 mM sodium phytate in 100 mM Tris HCl buffer pH 7.5, supplemented with

1mM $CaCl_2$ for 30 minutes at 37(C. After incubation the reaction is stopped by adding 750(l of 5% trichloroacetic acid. Phosphate released was measured against phosphate standard spectrophotometrically at 700nm after adding 1500(l of the color reagent (4 volumes of 1.5% ammonium molybdate in 5.5% sulfuric acid and 1 volume of 2.7% ferrous sulfate; Shimizu, M., 1992; Biosci. Biotech. Biochem., 56:1266-1269). One unit of enzyme activity is defined as the amount of enzyme required to liberate one (mol Pi per min under assay conditions. Specific activity can be expressed in units of enzyme activity per mg of protein.

The enzyme of the present invention has enzymatic activity with respect to the hydrolysis of phytate to inositol and free phosphate.

The polynucleotide of the present invention may be in the form of DNA which DNA includes cDNA, genomic DNA, and synthetic DNA. The DNA may be double-stranded or single-stranded, and if single stranded may be the coding strand or non-coding (anti-sense) strand. The coding sequence which encodes the mature enzyme may be identical to the coding sequences shown in Figure 1 and/or that of the deposited clone (SEQ ID NO:1), or may be a different coding sequence which coding sequence, as a result of the redundancy or degeneracy of the genetic code, encodes the same mature enzyme as the DNA of Figure 1 (*e.g.*, SEQ ID NO:1).

The polynucleotide which encodes for the mature enzyme of Figure 1 (*e.g.*, SEQ ID NO:2) may include, but is not limited to: only the coding sequence for the mature enzyme; the coding sequence for the mature enzyme and additional coding sequence such as a leader sequence or a proprotein sequence; the coding sequence for the mature enzyme (and optionally additional coding sequence) and non-coding sequence, such as introns or non-coding sequence 5' and/or 3' of the coding sequence for the mature enzyme.

Thus, the term "polynucleotide encoding an enzyme (protein)" encompasses a polynucleotide which includes only coding sequence for the enzyme as well as a polynucleotide which includes additional coding and/or non-coding sequence.

The present invention further relates to variants of the hereinabove described polynucleotides which encode for analogs and derivatives of the enzyme having the deduced amino acid sequence of Figure 1 (*e.g.*, SEQ ID NO:2). The variant of the polynucleotide may be a naturally occurring allelic variant of the polynucleotide or a non-naturally occurring variant of the polynucleotide.

Thus, the present invention includes polynucleotides encoding the same mature enzyme as shown in Figure 1 as well as variants of such polynucleotides which variants encode for a derivative or analog of the enzyme of Figure 1. Such nucleotide variants include deletion variants, substitution variants and addition or insertion variants.

As hereinabove indicated, the polynucleotide may have a coding sequence which is a naturally occurring allelic variant of the coding sequence shown in Figure 1. As known in the art, an allelic variant is an alternate form of a polynucleotide sequence which may have a substitution, deletion or addition of one or more nucleotides, which does not substantially alter the function of the encoded enzyme.

The present invention also includes polynucleotides, wherein the coding sequence for the mature enzyme may be fused in the same reading frame to a polynucleotide sequence which aids in expression and secretion of an enzyme from a host cell, for example, a leader sequence which functions to control transport of an enzyme from the cell. The enzyme having a leader sequence is a preprotein and may have the leader sequence cleaved by the host cell to form the mature form of the enzyme. The polynucleotides may also encode for a proprotein which is the mature protein plus additional 5' amino acid residues. A mature protein having a prosequence is a proprotein and is an inactive

form of the protein. Once the prosequence is cleaved an active mature protein remains.

Thus, for example, the polynucleotide of the present invention may encode for a mature enzyme, or for an enzyme having a prosequence or for an enzyme having both a prosequence and a presequence (leader sequence).

The present invention further relates to a enzyme which has the deduced amino acid sequence of Figure 1, as well as analogs and derivatives of such enzyme.

The terms "derivative" and "analog" when referring to the enzyme of Figure 1 means a enzyme which retains essentially the same biological function or activity as such enzyme. Thus, an analog includes a proprotein which can be activated by cleavage of the proprotein portion to produce an active mature enzyme.

The enzyme of the present invention may be a recombinant enzyme, a natural enzyme or a synthetic enzyme, preferably a recombinant enzyme.

The derivative or analog of the enzyme of Figure 1 may be (i) one in which one or more of the amino acid residues are substituted with an amino acid residue which is not encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature enzyme is fused with another compound, such as a compound to increase the half-life of the enzyme (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature enzyme, such as a leader or secretory sequence or a sequence which is employed for purification of the mature enzyme or a proprotein sequence. Such derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

The enzymes and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

The term "isolated" means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or enzyme present in a living animal is not isolated, but the same polynucleotide or enzyme, separated from some or all of the coexisting materials in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or enzymes could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment.

The enzymes of the present invention includes an enzyme of Figure 1 (in particular the mature enzyme) as well as enzymes which have at least 70% similarity (preferably at least 70% identity) to an enzyme of Figure 1 and more preferably at least 90% similarity (more preferably at least 90% identity) to an enzyme of Figure 1 and still more preferably at least 95% similarity (still more preferably at least 95% identity) to an enzyme of Figure 1 and also include portions of such enzymes with such portion of the enzyme generally containing at least 30 amino acids and more preferably at least 50 amino acids.

As known in the art "similarity" between two enzymes is determined by comparing the amino acid sequence and its conserved amino acid substitutes of one enzyme to the sequence of a second enzyme. Similarity in nucleic acid and amino acid sequences may be determined by procedures and algorithms which are well-known in the art. Such procedures and algorithms include, for example, a BLAST program (Basic Local Alignment Search Tool at the National Center for Biological Information), ALIGN, AMAS (Analysis of Multiply Aligned Sequences), AMPS (Protein Multiple Sequence Alignment), ASSET (Aligned Segment Statistical Evaluation Tool), BANDS, BESTSCOR, BIOSCAN (Biological Sequence Comparative Analysis Node), BLIMPS (BLocks IMProved Searcher), FASTA, Intervals & Points, BMB, CLUSTAL V,

CLUSTAL W, CONSENSUS, LCONSENSUS, WCONSENSUS, Smith-Waterman algorithm, DARWIN, Las Vegas algorithm, FNAT (Forced Nucleotide Alignment Tool), Framealign, Framesearch, DYNAMIC, FILTER, FSAP (Fristensky Sequence Analysis Package), GAP (Global Alignment Program), GENAL, GIBBS, GenQuest, ISSC (Sensitive Sequence Comparison), LALIGN (Local Sequence Alignment), LCP (Local Content Program), MACAW (Multiple Alignment Construction & Analysis Workbench), MAP (Multiple Alignment Program), MBLKP, MBLKN, PIMA (Pattern-Induced Multi-sequence Alignment), SAGA (Sequence Alignmnet by Genetic ALgorithm) and WHAT-IF.

A variant, i.e. a "analog" or "derivative" enzyme, and reference enzyme may differ in amino acid sequence by one or more substitutions, additions, deletions, fusions and truncations, which may be present in any combination.

Among preferred variants are those that vary from a reference by conservative amino acid substitutions. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics. Typically seen as conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr.

Most highly preferred are variants which retain the same biological function and activity as the reference polypeptide from which it varies.

Fragments or portions of the enzymes of the present invention may be employed for producing the corresponding full-length enzyme by peptide synthesis; therefore, the fragments may be employed as intermediates for producing the full-length enzymes. Fragments or portions of the polynucleotides of the present invention may be used to

synthesize full-length polynucleotides of the present invention.

The present invention also relates to vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of enzymes of the invention by recombinant techniques.

Host cells are genetically engineered (transduced or transformed or transfected) with the vectors containing the polynucleotides of this invention. Such vectors may be, for example, a cloning vector or an expression vector. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the genes of the present invention. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

The polynucleotides of the present invention may be employed for producing enzymes by recombinant techniques. Thus, for example, the polynucleotide may be included in any one of a variety of expression vectors for expressing an enzyme. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host.

The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art. Such procedures and others are deemed to be within the scope of those skilled in the art.

The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. As representative examples of such promoters, there may be mentioned: LTR or SV40 promoter, the *E. coli. lac* or *trp*, the phage lambda P_L promoter and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses. The expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in *E. coli*.

The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the protein.

As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as *E. coli*, Streptomyces, *Bacillus subtilis*; fungal cells, such as yeast; insect cells such as *Drosophila S2* and *Spodoptera Sf9*; animal cells such as CHO, COS or Bowes melanoma; adenoviruses; plant cells, *etc*. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

More particularly, the present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter,

operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example; Bacterial: pQE70, pQE60, pQE-9 (Qiagen), pBluescript II (Stratagene); pTRC99a, pKK223-3, pDR540, pRIT2T (Pharmacia); Eukaryotic: pXT1, pSG5 (Stratagene) pSVK3, pBPV, pMSG, pSVLSV40 (Pharmacia). However, any other plasmid or vector may be used as long as they are replicable and viable in the host.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda P_R, P_L and trp. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (Davis, L., Dibner, M., Battey, I., Basic Methods in Molecular Biology, (1986)).

The constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the enzymes of the invention can be synthetically produced by conventional peptide synthesizers.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells

under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), the disclosure of which is hereby incorporated by reference.

Transcription of the DNA encoding the enzymes of the present invention by higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act on a promoter to increase its transcription. Examples include the SV40 enhancer on the late side of the replication origin bp 100 to 270, a cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), Å-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated enzyme. Optionally, the heterologous sequence can encode a fusion enzyme including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation

and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli, Bacillus subtilis, Salmonella typhimurium* and various species within the genera Pseudomonas, Streptomyces, and Staphylococcus, although others may also be employed as a matter of choice.

As a representative but nonlimiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period.

Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, such methods are well known to those skilled in the art.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell, 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

The enzyme can be recovered and purified from recombinant cell cultures by methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

The enzymes of the present invention may be a naturally purified product, or a product of chemical synthetic procedures, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial, yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the enzymes of the present invention may be glycosylated or may be non-glycosylated. Enzymes of the invention may or may not also include an initial methionine amino acid residue.

The enzyme of this invention may be employed for any purpose in which such enzyme activity is necessary or desired. In a preferred embodiment the enzyme is employed for

catalyzing the hydrolysis of phytate. The degradation of phytate may be used in animal feed.

In a preferred embodiment, the enzyme of the present invention is a phytase enzyme which is stable to heat and is heat resistant and catalyzes the enzymatic hydrolysis of phytate, *i.e.*, the enzyme is able to renature and regain activity after a brief (*i.e.*, 5 to 30 seconds), or longer period, for example, minutes or hours, exposure to temperatures of 50(C optimum above 50(C.

The enzymes, their fragments or other derivatives, or analogs thereof, or cells expressing them can be used as an immunogen to produce antibodies thereto. These antibodies can be, for example, polyclonal or monoclonal antibodies. The present invention also includes chimeric, single chain, and humanized antibodies, as well as Fab fragments, or the product of an Fab expression library. Various procedures known in the art may be used for the production of such antibodies and fragments.

Antibodies generated against the enzymes corresponding to a sequence of the present invention can be obtained by direct injection of the enzymes into an animal or by administering the enzymes to an animal, preferably a nonhuman. The antibody so obtained will then bind the enzymes itself. In this manner, even a sequence encoding only a fragment of the enzymes can be used to generate antibodies binding the whole native enzymes. Such antibodies can then be used to isolate the enzyme from cells expressing that enzyme.

For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler and Milstein, 1975, Nature, 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., 1983, Immunology Today 4:72), and

the EBV-hybridoma technique to produce human monoclonal antibodies (Cole, et al., 1985, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96).

Techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce single chain antibodies to immunogenic enzyme products of this invention. Also, transgenic mice may be used to express humanized antibodies to immunogenic enzyme products of this invention.

Antibodies generated against the enzyme of the present invention may be used in screening for similar enzymes from other organisms and samples. Such screening techniques are known in the art. Antibodies may also be employed as a probe to screen gene libraries generated from this or other organisms to identify this or cross reactive activities.

Isolation and purification of polypeptides produced in the systems described above can be carried out using conventional methods, appropriate for the particular system. For example, preparative chromatography and immunological separations employing antibodies, such as monoclonal or polyclonal antibodies, can be used.

The term "antibody," as used herein, refers to intact immunoglobulin molecules, as well as fragments of immunoglobulin molecules, such as Fab, Fab', $(Fab')_2$, Fv, and SCA fragments, that are capable of binding to an epitope of an endoglucanase polypeptide. These antibody fragments, which retain some ability to selectively bind to the antigen (*e.g.*, an endoglucanase antigen) of the antibody from which they are derived, can be made using well known methods in the art (see, *e.g.*, Harlow and Lane, *supra*), and are described further, as follows.

(1) A Fab fragment consists of a monovalent antigen-binding fragment of an antibody molecule, and can be produced by digestion of a whole antibody molecule with the enzyme papain, to yield a fragment consisting of an intact light chain and a portion of a heavy chain.

(2) A Fab' fragment of an antibody molecule can be obtained by treating a whole antibody molecule with pepsin, followed by reduction, to yield a molecule consisting of an intact light chain and a portion of a heavy chain. Two Fab' fragments are obtained per antibody molecule treated in this manner.

(3) A $(Fab')_2$ fragment of an antibody can be obtained by treating a whole antibody molecule with the enzyme pepsin, without subsequent reduction. A $(Fab')_2$ fragment is a dimer of two Fab' fragments, held together by two disulfide bonds.

(4) An Fv fragment is defined as a genetically engineered fragment containing the variable region of a light chain and the variable region of a heavy chain expressed as two chains.

(5) A single chain antibody ("SCA") is a genetically engineered single chain molecule containing the variable region of a light chain and the variable region of a heavy chain, linked by a suitable, flexible polypeptide linker.

As used in this invention, the term "epitope" refers to an antigenic determinant on an antigen, such as a phytase polypeptide, to which the paratope of an antibody, such as an phytase-specific antibody, binds. Antigenic determinants usually consist of chemically active surface groupings of molecules, such as amino acids or sugar side chains, and can have specific three-dimensional structural characteristics, as well as specific charge characteristics.

As is mentioned above, antigens that can be used in producing phytase-specific antibodies include phytase polypeptides, *e.g.*, any of the phytase shown in Figures 1 polypeptide fragments. The polypeptide or peptide used to immunize an animal can be obtained by standard recombinant, chemical synthetic, or purification methods. As is well known in the art, in order to increase immunogenicity, an antigen can be conjugated to a carrier protein. Commonly used carriers include keyhole limpet hemocyanin (KLH), thyroglobulin, bovine serum albumin (BSA), and tetanus toxoid. The coupled peptide is then used to immunize the animal (*e.g.*, a mouse, a rat, or a rabbit). In addition to such carriers, well known adjuvants can be administered with the antigen to facilitate induction of a strong immune response.

Phytase-specific polyclonal and monoclonal antibodies can be purified, for example, by binding to, and elution from, a matrix containing a phytase polypeptide, *e.g.*, the phytase polypeptide (or fragment thereof) to which the antibodies were raised. Additional methods for antibody purification and concentration are well known in the art and can be practiced with the phytase-specific antibodies of the invention (see, for example, C oligan, *et al.*, Unit 9, *Current Protocols in Immunology*, Wiley Interscience, 1994).

Anti-idiotype antibodies corresponding to phytase-specific antigens are also included in the invention, and can be produced using standard methods. These antibodies are raised to phytase-specific antibodies, and thus mimic phytase-specific epitopes.

The members of a pair of molecules (*e.g.*, an antibody-antigen pair or a nucleic acid pair) are said to "specifically bind" to each other if they bind to each other with greater affinity than to other, non-specific molecules. For example, an antibody raised against an antigen to which it binds more efficiently than to a non-specific protein can be described as specifically binding to the antigen. (Similarly, a nucleic acid probe can be described as specifically binding to a nucleic acid target if it forms a specific duplex with the target by base pairing interactions (see above).)

The present invention is further described with reference to the following examples; however, it is to be understood that the present invention is not limited to such examples. All parts or amounts, unless otherwise specified, are by weight.

In one aspect of the invention, a method for producing an phytase enzyme, such as those shown in Figures 1, is provided. The method includes growing a host cell which contains a polynucleotide encoding the enzyme (*e.g.*, SEQ ID NO: 1), under conditions which allow the expression of the nucleic acid, and isolating the enzyme encoded by the nucleic acid. Methods of culturing the host cell are described in the Examples and are known by those of skill in the art.

In another embodiment, the invention provides a method for catalyzing the hydrolysis of phytate to inositol and free phosphate with release of minerals from the phytic acid complex. The method includes contacting phytate with a degrading effective amount of an enzyme of the invention, such as the enzyme shown in SEQ ID NO:1. The term "degrading effective" amount refers to the amount of enzyme which is required to degrade at least 50% of the phytate, as compared to phytate not contacted with the enzyme. Preferably, at least 80% of the phytate is degraded.

In another embodiment, the invention provides a method for hydrolyzing phospho-mono-ester bonds in phosphate, the method including administering an effective amount of an enzyme of the invention (*e.g.*, SEQ ID NO:1), to yield inositol and free phosphate. An "effective" amount refers to the amount of enzyme which is required to hydrolyze at least 50% of the phospho-mono-ester bonds, as compared to phytate not contacted with the enzyme. Preferably, at least 80% of the bonds are hydrolyzed.

In order to facilitate understanding of the following examples certain frequently occurring methods and/or terms will be described.

"Plasmids" are designated by a lower case p preceded and/or followed by capital letters and/or numbers. The starting plasmids herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids in accord with published procedures. In addition, equivalent plasmids to those described are known in the art and will be apparent to the ordinarily skilled artisan.

"Digestion" of DNA refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors and other requirements were used as would be known to the ordinarily skilled artisan. For analytical purposes, typically 1 ög of plasmid or DNA fragment is used with about 2 units of enzyme in about 20 öl of buffer solution. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 ög of DNA are digested with 20 to 250 units of enzyme in a larger volume. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation times of about 1 hour at 37°C are ordinarily used, but may vary in accordance with the supplier's instructions. After digestion the reaction is electrophoresed directly on a polyacrylamide gel to isolate the desired fragment.

Size separation of the cleaved fragments is generally performed using 8 percent polyacrylamide gel described by Goeddel, D. *et al.*, Nucleic Acids Res., 8:4057 (1980), for example.

"Oligonucleotides" refers to either a single stranded polydeoxynucleotide or two complementary polydeoxynucleotide strands which may be chemically synthesized. Such synthetic oligonucleotides may or may not have a 5' phosphate. Those that do not will not ligate to another oligonucleotide without adding a phosphate with an ATP in the presence of a kinase. A synthetic oligonucleotide will ligate to a fragment that has not been dephosphorylated.

"Ligation" refers to the process of forming phosphodiester bonds between two double

stranded nucleic acid fragments (Maniatis, T., et al., Id., p. 146). Unless otherwise provided, ligation may be accomplished using known buffers and conditions with 10 units of T4 DNA ligase ("ligase") per 0.5 (g of approximately equimolar amounts of the DNA fragments to be ligated.

Unless otherwise stated, transformation was performed as described in the method of Sambrook, Fritsch and Maniatus, 1989. The following examples are intended to illustrate, but not to limit, the invention. While the procedures described in the examples are typical of those that can be used to carry out certain aspects of the invention, other procedures known to those skilled in the art can also be used. The following materials and methods were used in carrying out the experiments described in the examples.

Example 1

Isolation, Bacterial Expression and Purification of Phytase

E.coli B genomic DNA was obtained from Sigma (Catalog # D-2001), St. Louis, New Jersey.

The following primers were used to PCR amplify the gene directly from the genomic DNA:

5' primer gtttctgaattcaaggaggaatttaaATGAAAGCGATCTTAATCCCATT

3' primer gtttctggatccTTACAAACTGCACGCCGGTAT

Pfu polymerase in the PCR reaction, and amplification was performed according to manufacturers protocol (Stratagene Cloning Systems, Inc., La Jolla, CA).

PCR product was purified and purified product and pQE60 vector (Qiagen) were both digested with EcoRI and BglII restriction endonucleases (New England Biolabs) according to manufacturers protocols. Overnight ligations were performed using

standard protocols to yield pQE60.

The amplified sequences were inserted in frame with the sequence encoding for the RBS. The ligation mixture was then used to transform the *E. coli* strain M15/pREP4 (Qiagen, Inc.) by electroporation. M15/pREP4 contains multiple copies of the plasmid pREP4, which expresses the lacI repressor and also confers kanamycin resistance (Kan^r). Plasmid DNA was isolated and confirmed by restriction analysis. Clones containing the desired constructs were grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture was used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells were grown to an optical density 600 ($O.D.^{600}$) of between 0.4 and 0.6. IPTG ("Isopropyl-B-D-thiogalacto pyranoside") was then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression. Cells were grown an extra 3 to 4 hours. Cells were then harvested by centrifugation.

The primer sequences set out above may also be employed to isolate the target gene from the deposited material by hybridization techniques described above.

Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, within the scope of the appended claims, the invention may be practiced otherwise than as particularly described. It is to be understood that, while the invention has been described with reference to the above detailed description, the foregoing description is intended to illustrate, but not to limit, the scope of the invention. Other aspects, advantages, and modifications of the invention are within the scope of the following claims. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety.

What Is Claimed Is:

1. Substantially pure phytase having an amino acid sequence selected from the group consisting of SEQ ID NO:2.

2. An isolated polynucleotide sequence encoding a phytase of claim 1.

3. An isolated polynucleotide selected from the group consisting of:
   a) SEQ ID NO:1;
   b) SEQ ID NO:1 wherein T can also be U;

4. The polynucleotide of claim 2, wherein the polynucleotide is isolated from a prokaryote.

5. An expression vector including the polynucleotide of claim 2.

6. The vector of claim 5, wherein the vector is a plasmid.

7. The vector of claim 5, wherein the vector is a virus-derived.

8. A host cell transformed with the vector of claim 5.

9. The host cell of claim 8, wherein the cell is prokaryotic.

10. Antibodies that bind to the polypeptide of claim 1.

11. The antibodies of claim 10, wherein the antibodies are polyclonal.

12. The antibodies of claim 10, wherein the antibodies are monoclonal.

13. A method for producing an enzyme comprising growing a host cell of claim 8 under conditions which allow the expression of the nucleic acid and isolating the enzyme encoded by the nucleic acid.

14. A method for degrading phytase comprising contacting phytate with a degrading effective amount of the enzyme of claim 1.

15. A method for hydrolyzing the phospho-mono-ester bond in phytate comprising contacting an effective amount of the enzyme of claim 1 with phytate to hydrolyze the bond.

16. An animal feed composition comprising a phytase.

17. The composition of claim 16, wherein the phytase has an amino acid sequence as set forth in SEQ ID NO:2.

18. An isolated polynucleotide having at least 70% identity to a member selected from the group consisting of:

a) a polynucleotide encoding an enzyme comprising an amino acid sequence as set forth in SEQ ID NO:2;

b) a polynucleotide which is complementary to the polynucleotide of a); and

c) a polynucleotide comprising at least 15 bases of the polynucleotide of a) or b).

19. An oligonucleotide probe that hybridizes to a nucleic acid target region corresponding to a region selected from the nucleic acid sequence set forth in SEQ ID NO:1

20. The probe of claim 19, wherein the probe is from 10 to about 150 nucleotides in length.

21. The probe of claim 19, wherein the probe is from 10 to about 100 nucleotides in length.

22. The probe of claim 19, wherein the probe is from 10 to about 50 nucleotides in length.

23. The probe of claim 19, wherein the probe is from 10 to about 30 nucleotides in length.

24. The probe of claim 19, wherein the probe is from 10 to about 15 nucleotides in length.

25. The oligonucleotide probe of claim 19, wherein the probe comprises a segment of 10 contiguous bases which is at least 70% complementary to a target sequence of 10 contiguous nucleotides present in the target region.

26. The probe of claim 19, wherein the probe is detectably labeled.

27. The probe of claim 26, wherein the detectable label is selected from the group consisting of a radioisotope, a fluorescent compound, a bioluminescent compound, a chemiluminescent compound, a metal chelator or an enzyme.

28. The antibody of claim 10, wherein the antibody is detectably labeled.

29. An enzyme selected from the group consisting of:

a) an enzyme comprising an amino acid sequence which is at least 70% identical to the amino acid sequence set forth in SEQ ID NO:2; and

b) an enzyme which comprises at least 30 consecutive amino acid residues homologous with an enzyme of a).

(SEQ ID NO:1-nucleotide sequence and SEQ ID NO:2-amino acid sequence)

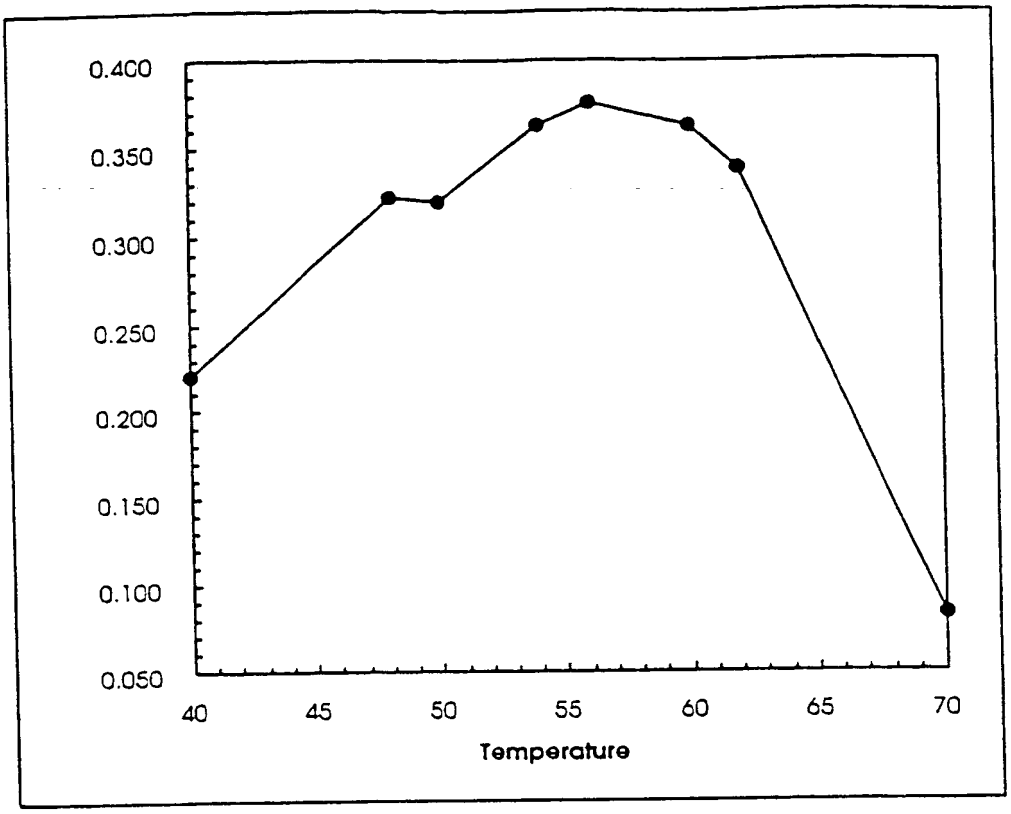
Escherichia coli B Phytase Sequence

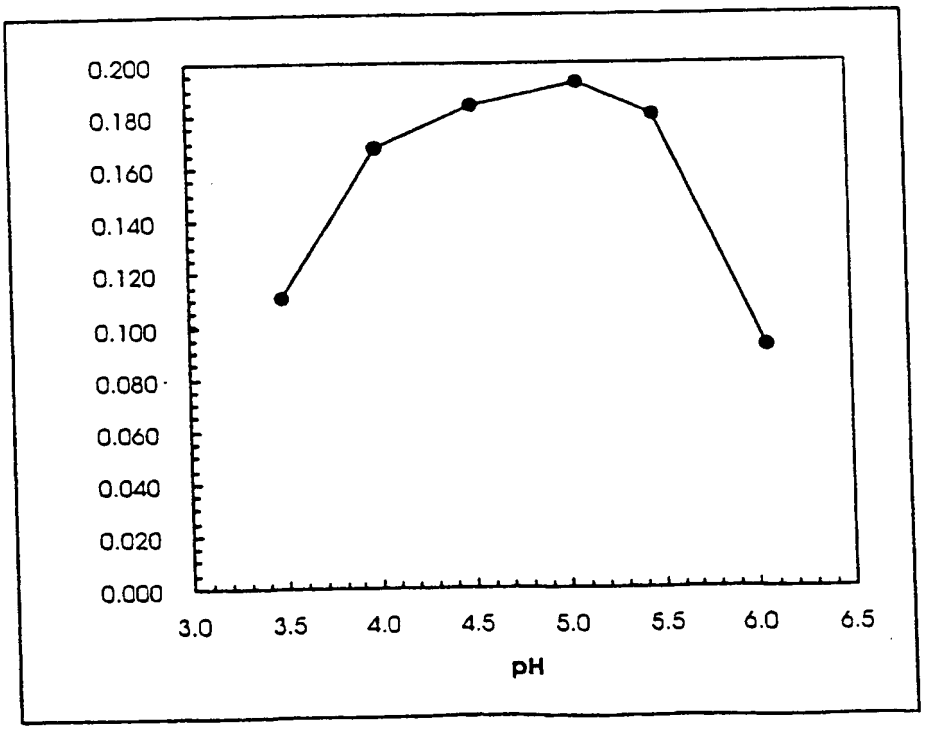
```
1
ATG AAA GCG ATC TTA ATC CCA TTT TTA TCT CTT CTG ATT CCG TTA ACC CCG
Met Lys Ala Ile Leu Ile Pro Phe Leu Ser Leu Leu Ile Pro Leu Thr Pro

CAA TCT GCA TTC GCT CAG AGT GAG CCG GAG CTG AAG CTG GAA AGT GTG GTG
Gln Ser Ala Phe Ala Gln Ser Glu Pro Glu Leu Lys Leu Glu Ser Val Val

ATT GTC AGT CGT CAT GGT GTG CGT GCT CCA ACC AAG GCC ACG CAA CTG ATG
Ile Val Ser Arg His Gly Val Arg Ala Pro Thr Lys Ala Thr Gln Leu Met

CAG GAT GTC ACC CCA GAC GCA TGG CCA ACC TGG CCG GTA AAA CTG GGT TGG
Gln Asp Val Thr Pro Asp Ala Trp Pro Thr Trp Pro Val Lys Leu Gly Trp

CTG ACA CCG CGN GGT GGT GAG CTA ATC GCC TAT CTC GGA CAT TAC CAA CGC
Leu Thr Pro Arg Gly Gly Glu Leu Ile Ala Tyr Leu Gly His Tyr Gln Arg

CAG CGT CTG GTA GCC GAC GGA TTG CTG GCG AAA AAG GGC TGC CCG CAG TCT
Gln Arg Leu Val Ala Asp Gly Leu Leu Ala Lys Lys Gly Cys Pro Gln Ser

GGT CAG GTC GCG ATT ATT GCT GAT GTC GAC GAG CGT ACC CGT AAA ACA GGC
Gly Gln Val Ala Ile Ile Ala Asp Val Asp Glu Arg Thr Arg Lys Thr Gly

GAA GCC TTC GCC GCC GGG CTG GCA CCT GAC TGT GCA ATA ACC GTA CAT ACC
Glu Ala Phe Ala Ala Gly Leu Ala Pro Asp Cys Ala Ile Thr Val His Thr

CAG GCA GAT ACG TCC AGT CCC GAT CCG TTA TTT AAT CCT CTA AAA ACT GGC
Gln Ala Asp Thr Ser Ser Pro Asp Pro Leu Phe Asn Pro Leu Lys Thr Gly

GTT TGC CAA CTG GAT AAC GCG AAC GTG ACT GAC GCG ATC CTC AGC AGG GCA
Val Cys Gln Leu Asp Asn Ala Asn Val Thr Asp Ala Ile Leu Ser Arg Ala

GGA GGG TCA ATT GCT GAC TTT ACC GGG CAT CGG CAA ACG GCG TTT CGC GAA
Gly Gly Ser Ile Ala Asp Phe Thr Gly His Arg Gln Thr Ala Phe Arg Glu

CTG GAA CGG GTG CTT AAT TTT CCG CAA TCA AAC TTG TGC CTT AAA CGT GAG
Leu Glu Arg Val Leu Asn Phe Pro Gln Ser Asn Leu Cys Leu Lys Arg Glu

AAA CAG GAC GAA AGC TGT TCA TTA ACG CAG GCA TTA CCA TCG GAA CTC AAG
Lys Gln Asp Glu Ser Cys Ser Leu Thr Gln Ala Leu Pro Ser Glu Leu Lys

GTG AGC GCC GAC AAT GTC TCA TTA ACC GGT GCG GTA AGC CTC GCA TCA ATG
Val Ser Ala Asp Asn Val Ser Leu Thr Gly Ala Val Ser Leu Ala Ser Met

CTG ACG GAG ATA TTT CTC CTG CAA CAA GCA CAG GGA ATG CCG GAG CCG GGG
Leu Thr Glu Ile Phe Leu Leu Gln Gln Ala Gln Gly Met Pro Glu Pro Gly

TGG GGA AGG ATC ACC GAT TCA CAC CAG TGG AAC ACC TTG CTA AGT TTG CAT
Trp Gly Arg Ile Thr Asp Ser His Gln Trp Asn Thr Leu Leu Ser Leu His

AAC GCG CAA TTT TAT TTG CTA CAA CGC ACG CCA GAG GTT GCC CGC AGC CGC
Asn Ala Gln Phe Tyr Leu Leu Gln Arg Thr Pro Glu Val Ala Arg Ser Arg

GCC ACC CCG TTA TTG GAT TTG ATC ATG GCA GCG TTG ACG CCC CAT CCA CCG
```

FIGURE 1 CONT'D

```
Ala Thr Pro Leu Leu Asp Leu Ile Met Ala Ala Leu Thr Pro His Pro Pro

CAA AAA CAG GCG TAT GGT GTG ACA TTA CCC ACT TCA GTA CTG TTT ATT GCC
Gln Lys Gln Ala Tyr Gly Val Thr Leu Pro Thr Ser Val Leu Phe Ile Ala

GGA CAC GAT ACT AAT CTG GCA AAT CTC GGC GGC GCA CTG GAG CTC AAC TGG
Gly His Asp Thr Asn Leu Ala Asn Leu Gly Gly Ala Leu Glu Leu Asn Trp

ACG CTT CCC GGT CAG CCG GAT AAC ACG CCG CCA GGT GGT GAA CTG GTG TTT
Thr Leu Pro Gly Gln Pro Asp Asn Thr Pro Pro Gly Gly Glu Leu Val Phe

GAA CGC TGG CGT CGG CTA AGC GAT AAC AGC CAG TGG ATT CAG GTT TCG CTG
Glu Arg Trp Arg Arg Leu Ser Asp Asn Ser Gln Trp Ile Gln Val Ser Leu

GTC TTC CAG ACT TTA CAG CAG ATG CGT GAT AAA ACG CCG CTG TCA TTA AAT
Val Phe Gln Thr Leu Gln Gln Met Arg Asp Lys Thr Pro Leu Ser Leu Asn

ACG CCG CCC GGA GAG GTG AAA CTG ACC CTG GCA GGA TGT GAA GAG CGA AAT
Thr Pro Pro Gly Glu Val Lys Leu Thr Leu Ala Gly Cys Glu Glu Arg Asn

GCG CAG GGC ATG TGT TCG TTG GCA GGT TTT ACG CAA ATC GTG AAT GAA GCA
Ala Gln Gly Met Cys Ser Leu Ala Gly Phe Thr Gln Ile Val Asn Glu Ala

CGC ATA CCG GCG TGC AGT TTG AGA TCT CAT CAC CAT CAC CAT CAC TAA  1323
Arg Ile Pro Ala Cys Ser Leu Arg Ser His His His His His His End
```

FIGURE 2
pH/Temperature Pr file and Stability





SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: KRETZ
(ii) TITLE OF INVENTION: NOVEL PHYTASE

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Fish & Richardson P.C.
(B) STREET: 4225 Executive Square, Suite 1400
(C) CITY: La Jolla
(D) STATE: California
(E) COUNTRY: US
(F) ZIP: 92037

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 08/910,798
(B) FILING DATE: August 13,1997
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: HAILE, PH.D., LISA A.
(B) REGISTRATION NUMBER: 38,347
(C) REFERENCE/DOCKET NUMBER: 09010/029001

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 619/678-5070
(B) TELEFAX: 619/678-5099

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1272 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:
(B) CLONE: PHYTASE

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
ATG AAA GCG ATC TTA ATC CCA TTT TTA TCT CTT CTG ATT CCG TTA ACC        48
Met Lys Ala Ile Leu Ile Pro Phe Leu Ser Leu Leu Ile Pro Leu Thr
  1               5                  10                  15

CCG CAA TCT GCA TTC GCT CAG AGT GAG CCG GAG CTG AAG CTG GAA AGT        96
Pro Gln Ser Ala Phe Ala Gln Ser Glu Pro Glu Leu Lys Leu Glu Ser
             20                  25                  30

GTG GTG ATT GTC AGT CGT CAT GGT GTG CGT GCT CCA ACC AAG GCC ACG       144
Val Val Ile Val Ser Arg His Gly Val Arg Ala Pro Thr Lys Ala Thr
         35                  40                  45

CAA CTG ATG CAG GAT GTC ACC CCA GAC GCA TGG CCA ACC TGG CCG GTA       192
Gln Leu Met Gln Asp Val Thr Pro Asp Ala Trp Pro Thr Trp Pro Val
     50                  55                  60

AAA CTG GGT TGG CTG ACA CCG CGN GGT GGT GAG CTA ATC GCC TAT CTC       240
Lys Leu Gly Trp Leu Thr Pro Arg Gly Gly Glu Leu Ile Ala Tyr Leu
 65                  70                  75                  80

GGA CAT TAC CAA CGC CAG CGT CTG GTA GCC GAC GGA TTG CTG GCG AAA       288
Gly His Tyr Gln Arg Gln Arg Leu Val Ala Asp Gly Leu Leu Ala Lys
                 85                  90                  95

AAG GGC TGC CCG CAG TCT GGT CAG GTC GCG ATT ATT GCT GAT GTC GAC       336
Lys Gly Cys Pro Gln Ser Gly Gln Val Ala Ile Ile Ala Asp Val Asp
            100                 105                 110

GAG CGT ACC CGT AAA ACA GGC CAG GCA GAT ACG TCC AGT CCC GAT CCG       384
Glu Arg Thr Arg Lys Thr Gly Gln Ala Asp Thr Ser Ser Pro Asp Pro
        115                 120                 125

TTA TTT AAT CCT CTA AAA ACT GGC GTT TGC CAA CTG GAT AAC GCG AAC       432
Leu Phe Asn Pro Leu Lys Thr Gly Val Cys Gln Leu Asp Asn Ala Asn
    130                 135                 140

GTG ACT GAC GCG ATC CTC AGC AGG GCA GGA GGG TCA ATT GCT GAC TTT       480
Val Thr Asp Ala Ile Leu Ser Arg Ala Gly Gly Ser Ile Ala Asp Phe
145                 150                 155                 160

ACC GGG CAT CGG CAA ACG GCG TTT CGC GAA CTG GAA CGG GTG CTT AAT       528
Thr Gly His Arg Gln Thr Ala Phe Arg Glu Leu Glu Arg Val Leu Asn
                165                 170                 175
```

```
TTT CCG CAA TCA AAC TTG TGC CTT AAA CGT GAG AAA CAG GAC GAA AGC        576
Phe Pro Gln Ser Asn Leu Cys Leu Lys Arg Glu Lys Gln Asp Glu Ser
            180                 185                 190

TGT TCA TTA ACG CAG GCA TTA CCA TCG GAA CTC AAG GTG AGC GCC GAC        624
Cys Ser Leu Thr Gln Ala Leu Pro Ser Glu Leu Lys Val Ser Ala Asp
        195                 200                 205

AAT GTC TCA TTA ACC GGT GCG GTA AGC CTC GCA TCA ATG CTG ACG GAG        672
Asn Val Ser Leu Thr Gly Ala Val Ser Leu Ala Ser Met Leu Thr Glu
    210                 215                 220

ATA TTT CTC CTG CAA CAA GCA CAG GGA ATG CCG GAG CCG GGG TGG GGA        720
Ile Phe Leu Leu Gln Gln Ala Gln Gly Met Pro Glu Pro Gly Trp Gly
225                 230                 235                 240

AGG ATC ACC GAT TCA CAC CAG TGG AAC ACC TTG CTA AGT TTG CAT AAC        768
Arg Ile Thr Asp Ser His Gln Trp Asn Thr Leu Leu Ser Leu His Asn
                245                 250                 255

GCG CAA TTT TAT TTG CTA CAA CGC ACG CCA GAG GTT GCC CGC AGC CGC        816
Ala Gln Phe Tyr Leu Leu Gln Arg Thr Pro Glu Val Ala Arg Ser Arg
            260                 265                 270

GCC ACC CCG TTA TTG GAT TTG ATC ATG GCA GCG TTG ACG CCC CAT CCA        864
Ala Thr Pro Leu Leu Asp Leu Ile Met Ala Ala Leu Thr Pro His Pro
        275                 280                 285

CCG CAA AAA CAG GCG TAT GGT GTG ACA TTA CCC ACT TCA GTA CTG TTT        912
Pro Gln Lys Gln Ala Tyr Gly Val Thr Leu Pro Thr Ser Val Leu Phe
    290                 295                 300

ATT GCC GGA CAC GAT ACT AAT CTG GCA AAT CTC GGC GGC GCA CTG GAG        960
Ile Ala Gly His Asp Thr Asn Leu Ala Asn Leu Gly Gly Ala Leu Glu
305                 310                 315                 320

CTC AAC TGG ACG CTT CCC GGT CAG CCG GAT AAC ACG CCG CCA GGT GGT       1008
Leu Asn Trp Thr Leu Pro Gly Gln Pro Asp Asn Thr Pro Pro Gly Gly
                325                 330                 335

GAA CTG GTG TTT GAA CGC TGG CGT CGG CTA AGC GAT AAC AGC CAG TGG       1056
Glu Leu Val Phe Glu Arg Trp Arg Arg Leu Ser Asp Asn Ser Gln Trp
            340                 345                 350

ATT CAG GTT TCG CTG GTC TTC CAG ACT TTA CAG CAG ATG CGT GAT AAA       1104
Ile Gln Val Ser Leu Val Phe Gln Thr Leu Gln Gln Met Arg Asp Lys
        355                 360                 365

ACG CCG CTG TCA TTA AAT ACG CCG CCC GGA GAG GTG AAA CTG ACC CTG       1152
Thr Pro Leu Ser Leu Asn Thr Pro Pro Gly Glu Val Lys Leu Thr Leu
    370                 375                 380

GCA GGA TGT GAA GAG CGA AAT GCG CAG GGC ATG TGT TCG TTG GCA GGT       1200
Ala Gly Cys Glu Glu Arg Asn Ala Gln Gly Met Cys Ser Leu Ala Gly
385                 390                 395                 400



TTT ACG CAA ATC GTG AAT GAA GCA CGC ATA CCG GCG TGC AGT TTG AGA       1248
Phe Thr Gln Ile Val Asn Glu Ala Arg Ile Pro Ala Cys Ser Leu Arg
                405                 410                 415

TCT CAT CAC CAT CAC CAT CAC TAA                                      1272
Ser His His His His His His
            420
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 423 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Lys Ala Ile Leu Ile Pro Phe Leu Ser Leu Leu Ile Pro Leu Thr
  1               5                  10                  15
Pro Gln Ser Ala Phe Ala Gln Ser Glu Pro Glu Leu Lys Leu Glu Ser
             20                  25                  30
Val Val Ile Val Ser Arg His Gly Val Arg Ala Pro Thr Lys Ala Thr
         35                  40                  45
Gln Leu Met Gln Asp Val Thr Pro Asp Ala Trp Pro Thr Trp Pro Val
     50                  55                  60
Lys Leu Gly Trp Leu Thr Pro Arg Gly Gly Glu Leu Ile Ala Tyr Leu
 65                  70                  75                  80
Gly His Tyr Gln Arg Gln Arg Leu Val Ala Asp Gly Leu Leu Ala Lys
                 85                  90                  95
Lys Gly Cys Pro Gln Ser Gly Gln Val Ala Ile Ile Ala Asp Val Asp
            100                 105                 110
Glu Arg Thr Arg Lys Thr Gly Gln Ala Asp Thr Ser Ser Pro Asp Pro
        115                 120                 125
Leu Phe Asn Pro Leu Lys Thr Gly Val Cys Gln Leu Asp Asn Ala Asn
    130                 135                 140
Val Thr Asp Ala Ile Leu Ser Arg Ala Gly Gly Ser Ile Ala Asp Phe
145                 150                 155                 160
Thr Gly His Arg Gln Thr Ala Phe Arg Glu Leu Glu Arg Val Leu Asn
                165                 170                 175
Phe Pro Gln Ser Asn Leu Cys Leu Lys Arg Glu Lys Gln Asp Glu Ser
            180                 185                 190
Cys Ser Leu Thr Gln Ala Leu Pro Ser Glu Leu Lys Val Ser Ala Asp
        195                 200                 205
Asn Val Ser Leu Thr Gly Ala Val Ser Leu Ala Ser Met Leu Thr Glu
    210                 215                 220
Ile Phe Leu Leu Gln Gln Ala Gln Gly Met Pro Glu Pro Gly Trp Gly
225                 230                 235                 240
Arg Ile Thr Asp Ser His Gln Trp Asn Thr Leu Leu Ser Leu His Asn
                245                 250                 255
Ala Gln Phe Tyr Leu Leu Gln Arg Thr Pro Glu Val Ala Arg Ser Arg
            260                 265                 270
Ala Thr Pro Leu Leu Asp Leu Ile Met Ala Ala Leu Thr Pro His Pro
        275                 280                 285
Pro Gln Lys Gln Ala Tyr Gly Val Thr Leu Pro Thr Ser Val Leu Phe
    290                 295                 300

Ile Ala Gly His Asp Thr Asn Leu Ala Asn Leu Gly Gly Ala Leu Glu
305                 310                 315                 320
Leu Asn Trp Thr Leu Pro Gly Gln Pro Asp Asn Thr Pro Pro Gly Gly
                325                 330                 335
Glu Leu Val Phe Glu Arg Trp Arg Arg Leu Ser Asp Asn Ser Gln Trp
            340                 345                 350
Ile Gln Val Ser Leu Val Phe Gln Thr Leu Gln Gln Met Arg Asp Lys
        355                 360                 365
Thr Pro Leu Ser Leu Asn Thr Pro Pro Gly Glu Val Lys Leu Thr Leu
    370                 375                 380
Ala Gly Cys Glu Glu Arg Asn Ala Gln Gly Met Cys Ser Leu Ala Gly
385                 390                 395                 400
Phe Thr Gln Ile Val Asn Glu Ala Arg Ile Pro Ala Cys Ser Leu Arg
                405                 410                 415
Ser His His His His His His
```

420

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 49 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GTTTCTGAAT TCAAGGAGGA ATTTAAATGA AAGCGATCTT AATCCCATT 49

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTTTCTGGAT CCTTACAAAC TGCACGCCGG TAT 33

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/16728

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :Please See Extra Sheet.
US CL :Please See Extra Sheet.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 426/53, 656; 435/6, 196, 325, 320.1, 252.3; 530/387.9, 388.26; 536/23.1, 23.2, 24.3

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	DASSA, J. et al. The Complete Nucleotide Sequence of the Escherichia coli Gene appA Reveals Significant Homology between pH 2.5 Acid Phosphatase and Glucose-1-Phosphatase. J. Bacteriol. September 1990, Vol. 172, No. 9, pages 5497-5500, especially page 5498.	18 ----- 10-12, 31
X	US 5,436,156 A (VAN GORCOM et al.) 25 July 1995, col. 5, lines 10-34.	16
X	Database GenBank on STN. DOEBELI, H. et al. 'Oligonucleotide sequence (adaptor 9) from patent EP0282042'. GenBank. January 1994. Accession No. A02249.	19-23, 25 ------- 26,27

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 13 OCTOBER 1998	Date of mailing of th international search report 23 OCT 1998
Name and mailing address f the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile N . (703) 305-3230	Authorized fficer PETER TUNG Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)★

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/16728

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6): A23B 7/10; A23J 1/00; C07K 16/12; C07H 15/11, 21/02, 21/04; C12N 9/16, 15/85, 15/63, 1/20; C12Q 1/68

A. CLASSIFICATION OF SUBJECT MATTER:
US CL : 426/53, 656; 435/6, 196, 325, 320.1, 252.3; 530/387.9, 388.26; 536/23.1, 23.2, 24.3

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

MasPar Search (DNA and Protein sequences), APS, MEDLINE, CAPLUS, CANCERLIT, BIOSIS, EMBASE, WPIDS, DISSABS
search terms: Keith Kretz, phytase, antibody, feed, E. coli, bacteria